

Protein Sequence Searches - February 2005

All of the sequence databases on ABSS have recently been updated.

- Please note that the curators of the UniProt database have purged some temporary accession numbers from the most recent version of UniProt. These sequences have been assigned new permanent accession numbers. The new UniProt record may not contain the previous temporary accession number.
- If you encounter an accession number from an older search run against UniProt (results file extension **.rup**) that can no longer be found in the database, the permanent record with the new accession number can be found by searching the old accession number in the UniProt Protein, Archive database (UniPARC) at:

<http://www.pir.uniprot.org/database/archive.shtml>

If you have any questions regarding this information or your results, please contact any STIC searcher.

When submitting sequence search results for scanning into IFW, please include a copy of this attachment to assist any future Examiners or members of the public who may encounter UniProt temporary accession numbers.

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 19:35:54 ; Search time 13.5116 Seconds
(without alignments)
49.847 Million cell updates/sec

Title: SEQ1

Perfect score: 54

Sequence: 1 gywgyw 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

PIR 79:*

1: piri:*

2: piri:*

3: piri:*

4: piri:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	83	2 E69903	hypothetical prote
2	54	100.0	187	2 G83047	hypothetical prote
3	49	90.7	273	2 E95268	probable ABC trans
4	45	83.3	68	2 A11956	hypothetical prote
5	45	83.3	271	2 F83188	phosphatidate cyti
6	45	83.3	271	2 JC4832	phosphatidate cyti
7	42	77.8	563	2 AH2975	hypothetical prote
8	42	77.8	563	2 C98307	onpA protein limpo
9	42	77.8	260	2 T41644	hypothetical trp-a
10	41	75.9	280	2 F81984	hypothetical prote
11	41	75.9	280	2 H81038	conserved hypotet
12	41	75.9	312	2 C70475	cytochrome c bioge
13	41	75.9	357	1 G62900	probable hexosyltr
14	41	75.9	400	1 F69142	probable ABC trans
15	41	75.9	536	2 G95389	cytochrome c-type
16	41	75.9	664	2 B75532	probable sodium-de
17	41	75.9	999	2 F72453	probable sodium-de
18	40	74.1	211	2 S08522	hypothetical prote
19	40	74.1	448	2 T23263	probable peptide/a
20	40	74.1	545	2 A84432	hypothetical prote
21	40	74.1	568	2 E96648	histidine transpor
22	40	74.1	585	2 C84432	histidine transpor
23	40	74.1	586	2 S46236	amino acid transpo
24	40	74.1	602	1 S38111	hypothetical prote
25	40	74.1	1039	2 T28905	sucrose alpha-gluc
26	40	74.1	1827	1 UUUU	sucrose alpha-gluc
27	40	74.1	1841	2 T10799	ig heavy chain pre
28	39	72.2	142	2 C34903	ammonia monooxygen
29	39	72.2	276	2 A49853	

30 39 72.2 294 2 S76046
31 39 72.2 369 2 A37316
32 39 72.2 529 2 T35966
33 39 72.2 826 2 E96720
34 39 72.2 1057 4 B47521
35 39 72.2 1622 2 D86428
36 39 72.2 1823 2 T01369
37 39 72.2 1870 2 C47521
38 39 72.2 2150 2 T08165
39 38 70.4 15 2 PH1366
40 38 70.4 120 2 A54256
41 38 70.4 243 2 AB1191
42 38 70.4 243 2 AB1549
43 38 70.4 351 2 JCS121
44 38 70.4 355 2 G83003
45 38 70.4 638 2 S51266

ALIGNMENTS

RESULT 1

E69903

hypothetical protein yodI - Bacillus subtilis

C/Species: Bacillus subtilis

C/Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004

C/Accession: E69903

R/Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berte

C.; Bron, S.; Brouillet, S.; Bruschi, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Ch

A.; Ehrlich, S.D.; Emerson, P.T.; Entian, K.D.; Errington, J.; Fabret, C.; Ferrari, E.

Nature 390, 249-256, 1997

A/Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Galle

iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hullo, M.F

Koester, P.; Koningstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois

A.; Muthers, Lauber, J.; Lazarevic, V.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Mauee

Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetell

Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadate, Y.; Sato, T.; Scanlon

A/Authors: Schleich, S.; Schroeter, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Sero

akeuchi, M.; Tanakoshi, A.; Tanaka, T.; Terpstra, P.; Tognoni, A.; Tosato, V.; Uchiyama

T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida,

A/Authors: Yoshikawa, H.F.; Zumstein, E.; Yoshikawa, H.; Danchin, A.

A/Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.

A/Reference number: A69580; MUID:98044033; PMID:9384377

A/Accession: E69903

A/Status: preliminary; nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-83 <KUN>

A/Cross-references: UNIPROT:O34654; GB:Z99114; GB:AL009126; NID:G2634230; PIDN:CAB13852

A/Experimental source: strain 168

C/Genetics:

A/Gene: yodI

Query Match 100.0%; Score 54; DB 2; Length 83;

Best Local Similarity 100.0%; Pred. No. 0.16;

Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYMGYYW 7

|||||

DB 53 GYMGYYW 59

RESULT 2

G83047

hypothetical protein PA4793 [imported] - Pseudomonas aeruginosa (strain PA01)

C/Species: Pseudomonas aeruginosa

C/Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004

C/Accession: G83047

R/Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; B

adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim

.; Lory, S.; Olson, M.V.

Nature 406, 959-964, 2000

A/Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path

A/Reference number: A82950; MUID:20437337; PMID:10984043

A;Accession: G83047
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-187 <STO>
 A;Cross-references: UNIPROT:Q9HV15; GB:AE004892; GB:AE004091; NID:G9951049; PIDN:AAG0817
 A;Experimental source: strain PA01
 C;Genetics:
 A;Gene: PA4793

Query Match 100.0%; Score 54; DB 2; Length 187;
 Best Local Similarity 100.0%; Pred. No. 0.33;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYGGYV 7
 |||||
 DB 109 GYGGYV 115

RESULT 3

E95268
 C;Species: Sinorhizobium meliloti
 C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004
 C;Accession: E95268
 R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bower, L.; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C.; Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001
 A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti
 A;Reference number: A95262; MUID:21396509; PMID:11481432
 A;Accession: E95268
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-273 <KUR>

A;Cross-references: UNIPROT:Q93021; GB:AE006469; PIDN:AAK64711.1; PID:G14523112; GSPDB:G83188
 A;Experimental source: strain 1021, megaplasmid pSymA
 R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, A.; Ampe, F.; Barloy-Hubler, P.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.; L.; Hyman, R.W.; Jones, T.
 Science 293, 668-672, 2001
 A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lellaure, heault, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.
 A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.
 A;Reference number: A96039; MUID:21368234; PMID:11474104
 A;Contents: annotation
 C;Genetics:
 A;Gene: Sma0106
 A;Genome: plasmid
 C;Superfamily: oligopeptide permease protein oppB

Query Match 90.7%; Score 49; DB 2; Length 273;
 Best Local Similarity 85.7%; Pred. No. 2.5;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYGGYV 7
 |||||
 DB 94 GYGGYV 100

RESULT 4

A11956
 C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C;Accession: A11956
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi, Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
 DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: A11956
 A;Status: preliminary
 A;Molecule type: DNA

A;Residues: 1-68 <KUR>
 A;Cross-references: UNIPROT:Q8YXK8; GB:BA000019; PIDN:BAE73161.1; PID:G17130551; GSPDB:G83188
 A;Experimental source: strain PCC 7120
 C;Genetics:
 A;Gene: asr1204

Query Match 83.3%; Score 45; DB 2; Length 68;
 Best Local Similarity 85.7%; Pred. No. 2.7;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYGGYV 7
 |||||
 DB 61 GRWGGYV 67

RESULT 5

F83188
 C;Species: Pseudomonas aeruginosa
 C;Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
 C;Accession: F83188
 R;Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warren, P.; Hickey, M.J.; B.; Adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Llin, L.; Lory, S.; Olson, M.V.
 Nature 406, 959-964, 2000
 A;Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic pathogen
 A;Reference number: A82950; MUID:20437337; PMID:10984043
 A;Accession: F83188
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-271 <STO>

A;Cross-references: UNIPROT:Q59640; GB:AE004785; GB:AE004091; NID:G9949809; PIDN:AAG070
 A;Experimental source: strain PA01
 C;Genetics:
 A;Gene: cdsA; PA3651
 C;Superfamily: phosphatidate cytidyltransferase

Query Match 83.3%; Score 45; DB 2; Length 271;
 Best Local Similarity 85.7%; Pred. No. 9.6;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYGGYV 7
 |||||
 DB 102 GYGGYV 108

RESULT 6

JC4832
 C;Species: Pseudomonas aeruginosa
 C;Date: 15-Aug-1996 #sequence_revision 15-Oct-1996 #text_change 09-Jul-2004
 C;Accession: JC4832
 R;Taguchi, K.; Fukutomi, H.; Kuroda, A.; Kato, J.; Ohtake, H.
 Gene 172, 165-166, 1996
 A;Title: Cloning of the Pseudomonas aeruginosa gene encoding CDP-diglyceride synthetase
 A;Reference number: JC4832; MUID:96257274; PMID:8654980
 A;Accession: JC4832
 A;Molecule type: DNA
 A;Residues: 1-271 <TAG>

A;Cross-references: UNIPROT:Q59640; DDBJ:D50811; NID:G1262331; PIDN:BA009437.1; PID:G1262331
 C;Comment: This enzyme catalyses the condensation of CTP and phosphatidic acid to form CDP-diglyceride
 C;Genetics:
 A;Gene: cds
 C;Superfamily: phosphatidate cytidyltransferase

C;Keywords: nucleotidyltransferase

Query Match 83.3%; Score 45; DB 2; Length 271;
 Best Local Similarity 85.7%; Pred. No. 9.6;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYGGYV 7
 |||||

Db 102 GYWGGRW 108

RESULT 7
AH2975
hypothetical protein Atu3409 [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
C;Species: Agrobacterium tumefaciens
C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
C;Accession: AH2975
R;Wood, D.W.; Setubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I.
erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutayavin, T.; Levy, R.; Li, M.; McClell
; Karp, P.; Romero, P.; Zhang, S.
Science 294, 2317-2323, 2001
A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
ster, E.W.
A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
A;Reference number: AB2577; MUID:21608550; PMID:11743193
A;Accession: AH2975
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-563 <KUR>
A;Cross-references: UNIPROT:Q8UAG5; GB:AE008689; PIDN:AA144222.1; PID:g17741803; GSPDB:G
A;Experimental source: strain C58 (Dupont)
C;Genetics:
A;Gene: Atu3409
A;Map position: linear chromosome

Query Match 77.8%; Score 42; DB 2; Length 563;
Best Local Similarity 71.4%; Pred. No. 52;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GYWGGRW 7
||| ||
Db 444 GYWSNYW 450

RESULT 8
C98307
oPhA protein [imported] - Agrobacterium tumefaciens (strain C58, Cereon)
C;Species: Agrobacterium tumefaciens
C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
C;Accession: C98307
R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Qurollo, B.; Goldman,
A.; Liu, F.; Wollam, C.; Allinger, D.; Doughty, M.; Scott, C.; Lappas, C.; Markelz, B.;
Science 294, 2323-2328, 2001
A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
A;Reference number: A97359; MUID:21608551; PMID:11743194
A;Accession: C98307
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-563 <KUR>
A;Cross-references: UNIPROT:Q8UAG5; GB:AE007870; PIDN:AAK89981.1; PID:g15159946; GSPDB:G
C;Genetics:
A;Gene: AGP_L_2824
A;Map position: linear chromosome

Query Match 77.8%; Score 42; DB 2; Length 563;
Best Local Similarity 71.4%; Pred. No. 52;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 GYWGGRW 7
||| ||
Db 444 GYWSNYW 450

RESULT 9
T41644
hypothetical trp-asp repeat protein - fission yeast (Schizosaccharomyces pombe)
C;Species: Schizosaccharomyces pombe
C;Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 16-Aug-2004
C;Accession: T41644
R;Wood, V.; Rajandream, M.A.; Barrell, B.G.; Lucas, M.; Gaillardin, C.
submitted to the EMBL Data Library, January 1999

A;Reference number: Z22006
A;Accession: T41644
A;Status: preliminary; translated from GB/EMBL/DDAU
A;Molecule type: DNA
A;Residues: 1-760 <WOO>
A;Cross-references: UNIPROT:O94533; EMBL:AL035247; PIDN:CAA22842.1; GSPDB:GN00068; SPDB
A;Experimental source: strain 972h-; cosmid c895
C;Genetics:
A;Gene: SPDB:SPCC895.06
A;Map position: 3
A;Introns: 39/3; 649/3
C;Superfamily: WD repeat homology

Query Match 77.8%; Score 42; DB 2; Length 760;
Best Local Similarity 71.4%; Pred. No. 68;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYWGGRW 7
||| ||
Db 328 GFWGLW 334

RESULT 10
F81984
hypothetical protein NMA0646 [imported] - Neisseria meningitidis (strain Z2491 serogrou
C;Species: Neisseria meningitidis
C;Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C;Accession: F81984
R;Parkhill, J.; Achtman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; More
; Holtroyd, S.; Jagers, K.; Leather, S.; Moule, S.; Mungall, K.; Quail, M.A.; Rajandream
Nature 404, 502-506, 2000
A;Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A;Reference number: A81775; MUID:20222556; PMID:10761919
A;Accession: F81984
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-280 <PAR>
A;Cross-references: UNIPROT:Q9JVM6; GB:AL162753; GB:AL157959; NID:g7379120; PIDN:CA8839
A;Experimental source: serogroup A, strain Z2491
C;Genetics:
A;Gene: NMA0646
C;Superfamily: hypothetical protein HI1037

Query Match 75.9%; Score 41; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 3 WGGYW 7
||| ||
Db 99 WGGYW 103

RESULT 11
H81038
conserved hypothetical protein NMB1815 [imported] - Neisseria meningitidis (strain MC58
C;Species: Neisseria meningitidis
C;Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
C;Accession: H81038
R;Tettelin, H.; Saunders, N.J.; Heidelberg, J.; Jeffries, A.C.; Nelson, K.E.; Eisen, J.
Hickey, E.K.; Haft, D.H.; Salzberg, S.L.; White, O.; Fleischmann, R.D.; Dougherty, B.A.
ri, H.; Qin, H.; Vamathevan, J.; Gill, J.; Scarlato, V.; Maignani, V.; Pizza, M.
Science 287, 1809-1815, 2000
A;Authors: Grandi, G.; Sun, L.; Smith, H.O.; Fraser, C.M.; Moxon, E.R.; Rappuoli, R.; V
A;Title: Complete genome sequence of Neisseria meningitidis serogroup B strain MC58.
A;Reference number: AB1000; MUID:20175755; PMID:10710307
A;Accession: H81038
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-280 <TET>
A;Cross-references: UNIPROT:Q9JY00; GB:AE002531; GB:AE002098; NID:g7227065; PIDN:AAF421
A;Experimental source: serogroup B, strain MC58
C;Genetics:
A;Gene: NMB1815

C;Superfamily: hypothetical protein HI1037

Query Match 75.9%; Score 41; DB 2; Length 280;
Best Local Similarity 100.0%; Pred. No. 38;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WGGYW 7
|||
Db 99 WGGYW 103

RESULT 12

C70475

Cytochrome c biogenesis protein - Aquifex aeolicus

C;Species: Aquifex aeolicus

C;Date: 08-May-1998 #sequence_revision 08-May-1998 #text_change 09-Jul-2004

C;Accession: C70475

R;Deckert, G.; Warren, P.V.; Gaasterland, T.; Young, W.G.; Lenox, A.L.; Graham, D.E.; Ov

V. Nature 392, 353-358, 1998

A;Title: The complete genome of the hyperthermophilic bacterium Aquifex aeolicus.

A;Reference number: A70300; MUID:98196666; PMID:9537320

A;Accession: C70475

A;Status: preliminary; nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-312 <AQF>

A;Cross-references: UNIPROT:O67831; GB:AE000769; NID:q2984262; PIDN:AAO07795.1; PID:q298

A;Experimental source: strain VF5

C;Genetics:

A;Gene: hemX2

C;Superfamily: cytochrome c-type synthesis protein

Query Match 75.9%; Score 41; DB 2; Length 312;

Best Local Similarity 100.0%; Pred. No. 42;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WGGYW 7

|||

Db 244 WGGYW 248

RESULT 13

G69290

probable hexosyltransferase (EC 2.4.1.-) AF0327 [similarity] - Archaeoglobus fulgidus

C;Species: Archaeoglobus fulgidus

C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004

C;Accession: G69290

R;Klenk, H.P.; Clayton, R.A.; Tomb, J.F.; White, O.; Nelson, K.E.; Ketchum, K.A.; Dodson

; Fleischmann, R.D.; Quackenbush, J.; Lee, N.H.; Sutton, G.G.; Gill, S.; Kirkness, E.F.

Glodek, A.; Zhou, L.; Overbeek, R.; Gocayne, J.D.; Weidman, J.P.; McDonald, L.

Nature 390, 364-370, 1997

A;Authors: Utterback, T.; Cotton, M.D.; Spriggs, T.; Artiach, P.; Kaine, B.P.; Sykes, S.

Smith, H.O.; Woese, C.R.; Venter, J.C.

A;Title: The complete genome sequence of the hyperthermophilic, sulfate-reducing archae

A;Reference number: A69250; MUID:98049343; PMID:9389475

A;Accession: G69290

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-357 <KLE>

A;Cross-references: UNIPROT:O39920; GB:AE001082; GB:AE000782; NID:g2689405; PIDN:AA9090

C;Superfamily: probable hexosyltransferase ytxN

C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 75.9%; Score 41; DB 1; Length 357;

Best Local Similarity 83.3%; Pred. No. 48;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YWGGYW 7

|||

Db 120 YWGGYW 125

RESULT 14

F69142

probable hexosyltransferase (EC 2.4.1.-) WTH332 [similarity] - Methanobacterium thermoa

N;Alternate names: LPS biosynthesis rfbU related protein

C;Species: Methanobacterium thermoautotrophicum

C;Date: 20-Apr-2000 #sequence_revision 20-Apr-2000 #text_change 09-Jul-2004

C;Accession: F69142

R;Smith, D.R.; Doucette-Stamm, L.A.; Deloughery, C.; Lee, H.; Dubois, J.; Aldredge, T.;

Qiu, D.; Spadafora, R.; Viceire, R.; Wang, Y.; Wierzbowski, J.; Gibson, R.; Jiwani, N.

J. Bacteriol. 179, 7135-7155, 1997

A;Title: Complete genome sequence of Methanobacterium thermoautotrophicum Delta H: func

A;Reference number: A69000; MUID:98037514; PMID:9371463

A;Accession: F69142

A;Status: nucleic acid sequence not shown; translation not shown

A;Molecule type: DNA

A;Residues: 1-400 <MTH>

A;Cross-references: UNIPROT:O26432; GB:AE000818; GB:AE000666; NID:g2621384; PIDN:AA84843

A;Experimental source: strain Delta H

C;Genetics:

A;Gene: WTH332

C;Superfamily: probable hexosyltransferase ytxN

C;Keywords: glycosyltransferase; hexosyltransferase

Query Match 75.9%; Score 41; DB 1; Length 400;

Best Local Similarity 83.3%; Pred. No. 53;

Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YWGGYW 7

|||

Db 155 YWGGYW 160

RESULT 15

G95389

probable ABC transporter, periplasmic solute-binding protein Sma1860 [imported] - Sinorh

C;Species: Sinorhizobium meliloti

C;Date: 24-Aug-2001 #sequence_revision 24-Aug-2001 #text_change 09-Jul-2004

C;Accession: G95389

R;Barnett, M.J.; Fisher, R.F.; Jones, T.; Komp, C.; Abola, A.P.; Barloy-Hubler, F.; Bow

; Kalman, S.; Keating, D.H.; Palm, C.; Peck, M.C.; Surzycki, R.; Wells, D.H.; Yeh, K.C

Proc. Natl. Acad. Sci. U.S.A. 98, 9883-9888, 2001

A;Title: Nucleotide sequence and predicted functions of the entire Sinorhizobium meliloti

A;Reference number: A95262; MUID:21396509; PMID:11481432

A;Accession: G95389

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-536 <KUR>

A;Cross-references: UNIPROT:Q92Y64; GB:AE006469; PIDN:AAK65681.1; PID:g14524171; GSPDB:(

A;Experimental source: strain 1021, megaplasmid pSyma

R;Galibert, F.; Finan, T.M.; Long, S.R.; Puhler, A.; Abola, P.; Ampe, F.; Barloy-Hubler,

Pela, D.; Chain, P.; Cowie, A.; Davis, R.W.; Dreano, S.; Federspiel, N.A.; Fisher, R.F.;

L.; Hyman, R.W.; Jones, T.

Science 293, 668-672, 2001

A;Authors: Kahn, D.; Kahn, M.L.; Kalman, S.; Keating, D.H.; Kiss, E.; Komp, C.; Lelaure,

hebaull, P.; Vandenbol, M.; Vorholter, F.J.; Weidner, S.; Wells, D.H.; Wong, K.; Yeh, K.C

A;Title: The composite genome of the legume symbiont Sinorhizobium meliloti.

A;Reference number: A96039; MUID:21368234; PMID:11474104

A;Contents: annotation

C;Genetics:

A;Gene: Sma1860

A;Genome: plasmid

Query Match 75.9%; Score 41; DB 2; Length 536;

Best Local Similarity 71.4%; Pred. No. 69;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYGGYW 7

|||

Db 417 GYGGYW 423

Search completed: April 18, 2005, 19:51:53

Job time : 15.5116 secs

This Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 19:37:28 ; Search time 62.3488 Seconds
(without alignments)
57.492 Million cell updates/sec

Title: SEQ1

Perfect score: 54

Sequence: 1 gywgggyw 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : UniProt_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	83	1 YODI_BACSU	Q34654 bacillus su
2	54	100.0	187	2 Q9HV15	Q9HV15 pseudomonas
3	54	100.0	189	2 Q88QB8	Q88QB8 pseudomonas
4	54	100.0	440	2 Q64VU6	Q64VU6 bacteroides
5	49	90.7	273	2 Q93021	Q93021 rhizobium m
6	49	90.7	273	2 Q89KL1	Q89KL1 bradyrhizob
7	48	88.9	296	2 Q82LB6	Q82LB6 streptomyce
8	47	87.0	363	2 Q8U043	Q8U043 pyrococcus
9	46	85.2	214	2 Q8A8U4	Q8A8U4 bacteroides
10	45	83.3	68	2 Q8YXK8	Q8YXK8 anabaena sp
11	45	83.3	106	2 Q9VZB1	Q9VZB1 drosophila
12	45	83.3	271	1 CDSA_PSEAE	Q59640 p phosphati
13	45	83.3	762	2 Q7SD72	Q7SD72 neurospora
14	44	81.5	302	2 Q6LS88	Q6LS88 photobacter
15	42	77.8	210	2 Q88RD4	Q88RD4 pseudomonas
16	42	77.8	261	1 CAPA_ASHGO	Q75d54 ashbya goss
17	42	77.8	273	2 Q81A84	Q81A84 bacillus ce
18	42	77.8	326	2 Q81A83	Q81A83 bacillus ce
19	42	77.8	368	2 Q73TG1	Q73TG1 mycobacteri
20	42	77.8	441	2 Q7N8B6	Q7N8B6 photorhabdu
21	42	77.8	563	2 Q8UAG5	Q8UAG5 agrobacteri
22	42	77.8	760	2 Q94533	Q94533 schizosacch
23	42	77.8	982	2 Q64Y78	Q64Y78 bacteroides
24	42	77.8	982	2 Q8A124	Q8A124 bacteroides
25	41.5	76.9	429	2 Q8AAZ7	Q8AAZ7 bacteroides
26	41.5	76.9	462	1 NORM_BURVI	Q9f5n7 burkholderi
27	41.5	76.9	468	2 Q62LW6	Q62LW6 burkholderi
28	41.5	76.9	468	2 Q63S57	Q63S57 burkholderi
29	41	75.9	236	2 Q63RF2	Q63RF2 burkholderi
30	41	75.9	280	2 Q5O469	Q5O469 neisseria g
31	41	75.9	280	2 Q9JWV6	Q9JWV6 neisseria m

32 41 75.9 280 2 Q9JY00 Q9JY00 neisseria m
33 41 75.9 286 2 Q62KZ9 Q62KZ9 burkholderi
34 41 75.9 312 2 Q67831 Q67831 aquifex aeo
35 41 75.9 357 2 Q29920 Q29920 archaeoglob
36 41 75.9 381 2 Q7NPV4 Q7NPV4 chromobacte
37 41 75.9 395 2 Q8XV46 Q8XV46 raiistonia s
38 41 75.9 395 2 Q9ADK5 Q9ADK5 streptomyce
39 41 75.9 396 2 Q62GN8 Q62GN8 burkholderi
40 41 75.9 396 2 Q63Q45 Q63Q45 burkholderi
41 41 75.9 400 2 Q26432 Q26432 methanobact
42 41 75.9 412 2 Q8TL58 Q8TL58 methanosarc
43 41 75.9 505 2 Q98718 Q98718 rhizobium l
44 41 75.9 536 2 Q6W118 Q6W118 rhizobium s
45 41 75.9 536 2 Q92Y64 Q92Y64 rhizobium m

ALIGNMENTS

RESULT 1

ID YODI_BACSU STANDARD; PRT; 83 AA.
AC Q34654;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 25-OCT-2004 (Rel. 45, Last annotation update)
DE Hypothetical protein yodi.
GN Name=yodi; Synonyms=yola; OrderedLocusNames=BSU19610;
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RA Wambutt R., Wedler H., Lapidus A., Sorokin A., Ehrlich S.D.;
RT "Sequence analysis of the Bacillus subtilis chromosome region between
RT the odhA and aspc loci cloned in a yeast artificial chromosome.";
RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA Chim S.-Y., Jeong Y.-M., Choi S.-K., Park S.-H.;
RT "Sequence analysis of the 30 kb region (182') of the Bacillus subtilis
RT chromosome containing the cge cluster";
RN [3]
RP Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RA MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriass R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruscia C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Conner I.F., Cummings N.J., Daniel R.A.,
RA Denist F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.F.,
RA Entian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.-Y., Glaser P., Goffeau A., Goldlighty E.J., Grandi G.,
RA Guiseppe G., Guy B.J., Haga K., Haeck J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,
RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
RA Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S.,
RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C.,
RA Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S.,
RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
RA Prescott A.M., Prescan E., Pujic P., Purnelle B., Rapoport G.,
RA Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B.,
RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
RA Scoffone F., Skiguchi J., Sekowska A., Seror S.J., Seror P.,
RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
RA Takenaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpetra P.,
RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,

```

RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
RA Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K.,
RA Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,
RA Yoshikawa H., Danchin A.;
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RT subtilis."
RL Nature 390:249-256(1997).
CC -----
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AF015775; AAB72056.1; -.
DR EMBL; AF006665; AAB81166.1; -.
DR EMBL; Z99114; CAB13852.1; -.
DR PIR; E69903; E69903.
DR Subtilisin; EGI13537; yodI.
DR InterPro; IPR008991; Transl_SH3_like.
DR Complete proteome; Hypothetical protein; Transmembrane.
FT TRANSGEM 58 80 Potential.
SQ SEQUENCE 83 AA; 9194 MW; 99F58EA2F0F36A43 CRC64;
Query Match 100.0%; Score 54; DB 1; Length 83;
Best Local Similarity 100.0%; Pred. No. 1.2;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GYWGYYW 7
Db 53 GYWGYYW 59
RESULT 2
Q9HV15 PRELIMINARY; PRT; 187 AA.
AC Q9HV15
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=PA4793;
OS Pseudomonas aeruginosa.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=287;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 15692 / PA01;
RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
RA Slover C.K., Pham X.-O.T., Erwin A.L., Mizoguchi S.D., Warrenner P.,
RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
RA Garber R.L., Goltzy L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
RT opportunistic pathogen."
RL Nature 406:959-964(2000).
DR EMBL; AE004892; AAG08179.1; -.
DR PIR; G83047; G83047.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 187 AA; 21281 MW; 8908E9EEEC51897 CRC64;
Query Match 100.0%; Score 54; DB 2; Length 187;
Best Local Similarity 100.0%; Pred. No. 2.4;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GYWGYYW 7
Db 109 GYWGYYW 115

```

```

RESULT 3
Q88QB8 PRELIMINARY; PRT; 189 AA.
ID Q88QB8
AC Q88QB8;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Lipoprotein, putative.
GN OrderedLocusNames=PP0576;
OS Pseudomonas putida (strain KT2440).
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
OC Pseudomonadaceae; Pseudomonas.
OX NCBI_TaxID=160488;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=22423060; PubMed=12534463;
RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
RA Moazed A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
RA Kiewitz C., Eisen J.A., Timmis K.N., Duesterhoeft A., Tuenmler B.,
RA Fraser C.M.;
RT "Complete genome sequence and comparative analysis of the
RT metabolically versatile Pseudomonas putida KT2440."
RL Environ. Microbiol. 4:799-808(2002).
DR EMBL; AE016776; AAN66203.1; -.
DR TIGR; PP0576; -.
KW Complete proteome; Lipoprotein.
SQ SEQUENCE 189 AA; 21449 MW; F0501C66D4EF483C CRC64;
Query Match 100.0%; Score 54; DB 2; Length 189;
Best Local Similarity 100.0%; Pred. No. 2.5;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GYWGYYW 7
Db 112 GYWGYYW 118
RESULT 4
Q64VU6 PRELIMINARY; PRT; 440 AA.
ID Q64VU6
AC Q64VU6;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Hypothetical protein.
GN ORFNames=BF1632;
OS Bacteroides fragilis.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=817;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YCH46;
RA Kuwahara T., Yamashita A., Hirakawa H., Nakayama H., Toh H., Okada N.,
RA Kuwaha S., Hattori M., Hayashi T., Ohnishi Y.;
RT "Genomic analysis of Bacteroides fragilis reveals extensive DNA
RT inversions regulating cell surface adaptation."
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2004).
DR EMBL; AP006841; BAD48380.1; -.
KW Hypothetical protein.
SQ SEQUENCE 440 AA; 49713 MW; 2EC38AFB04138439 CRC64;
Query Match 100.0%; Score 54; DB 2; Length 440;
Best Local Similarity 100.0%; Pred. No. 5.1;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY 1 GYGGGYW 7
Db 212 GYGGGYW 218

RESULT 5
ID Q93021 PRELIMINARY; PRT; 273 AA.
AC Q93021;
DT 01-DRC-2001 (T-EMBLrel. 19, Created)
DT 01-DRC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Putative ABC transporter, permease.
GN ORFNames=Swao106;
OS Rhizobium meliloti (Sinorhizobium meliloti).
OG Plasmid pSymA.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Sinorhizobium/Ensifer group; Sinorhizobium.
OX NCBI_TaxID=382;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=1021;
RX MEDLINE=21396509; PubMed=11481432; DOI=10.1073/pnas.161294798;
RA Barnett M.J., Fisher R.F., Jones T., Komp C., Abola A.P.,
RA Barloy-Hubier F., Bowser L., Capela D., Galibert F., Gozy J.,
RA Gurjal M., Hong A., Huizar L., Hyman R.W., Kahn D., Kahn M.L.,
RA Kalman S., Keating D.H., Palm C., Peck M.C., Surzycki R., Wells D.H.,
RA Yeh K.-C., Davis R.W., Federspiel N.A., Long S.R.;
RT "Nucleotide sequence and predicted functions of the entire
RT Sinorhizobium meliloti pSymA megaplasmid.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:9883-9888(2001).
CC -!- FUNCTION: Part of a binding-protein-dependent transport system.
CC Probably responsible for the translocation of the substrate across
CC the membrane (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the binding-protein-dependent transport
CC system permease family.
DR EMBL; AE007199; AAK64711.1; -.
DR FIR; E95268; E95268.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.
DR Pfam; PF00528; BPD transp.1; 1.
DR PROSITE; PS0928; ABC_TM1; 1.
KW Complete proteome; Plasmid; Transmembrane; Transport.
SQ SEQUENCE 273 AA; 29288 MW; 62884BA92EE49AD8 CRC64;

Query Match 90.7%; Score 49; DB 2; Length 273;
Best Local Similarity 85.7%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYGGGYW 7
Db 94 GYGGGYW 100

RESULT 6
ID Q89KL1 PRELIMINARY; PRT; 273 AA.
AC Q89KL1;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE ABC transporter permease protein.
GN OrderedLocusNames=bll4894;
OS Bradyrhizobium japonicum.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Bradyrhizobiaceae; Bradyrhizobium.
OX NCBI_TaxID=375;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=USDA110;
RX MEDLINE=22484998; PubMed=12597275;
RA Kaneko T., Nakamura Y., Sato S., Minamisawa K., Uchiumi T.,
RA Sasamoto S., Watanabe A., Idesawa K., Iriguchi M., Kawashima K.,
RA Kohara M., Matsuoka M., Shimpo S., Teurouka H., Wada T., Yamada M.,
RA Tabata S.;
RT "Complete genomic sequence of nitrogen-fixing symbiotic bacterium
RT Bradyrhizobium japonicum USDA110.";
RL DNA Res. 9:189-197(2002).
CC -!- FUNCTION: Part of a binding-protein-dependent transport system.
CC Probably responsible for the translocation of the substrate across
CC the membrane (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the binding-protein-dependent transport
CC system permease family.
DR EMBL; AP005953; BAC50159.1; -.
DR GO; GO:0016021; C:integral to membrane; IEA.
DR GO; GO:0005215; F:transporter activity; IEA.
DR GO; GO:0006810; P:transport; IEA.
DR InterPro; IPR000515; BPD transp.
DR Pfam; PF00528; BPD transp.1; 1.
DR PROSITE; PS0928; ABC_TM1; 1.
KW Complete proteome; Transmembrane; Transport.
SQ SEQUENCE 273 AA; 29042 MW; 6E3D8EE197497613 CRC64;

Query Match 90.7%; Score 49; DB 2; Length 273;
Best Local Similarity 85.7%; Pred. No. 17;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYGGGYW 7
Db 94 GYGGGYW 100

RESULT 7
ID Q82LB6 PRELIMINARY; PRT; 296 AA.
AC Q82LB6;
DT 01-JUN-2003 (T-EMBLrel. 24, Created)
DT 01-JUN-2003 (T-EMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (T-EMBLrel. 26, Last annotation update)
DE Putative sugar transporter.
GN OrderedLocusNames=SAV2094;
OS Streptomyces avermitilis.
OC Bacteria; Actinobacteria; Actinobacteridae; Actinomycetales;
OC Streptomycineae; Streptomycetaceae; Streptomyces.
OX NCBI_TaxID=33903;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=21477403; PubMed=11572948; DOI=10.1073/pnas.211433198;
RA Omura S., Ikeda H., Ishikawa J., Hanamoto A., Takahashi C.,
RA Shinose M., Takahashi Y., Horikawa H., Nakazawa H., Osonoe T.,
RA Kikuchi H., Shiba T., Sakaki Y., Hattori M.;
RT "Genome sequence of an industrial microorganism Streptomyces
RT avermitilis: deducing the ability of producing secondary
RT metabolites.";
RL Proc. Natl. Acad. Sci. U.S.A. 98:12215-12220(2001).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=MA-4680;
RX MEDLINE=22608306; PubMed=12692562;
RA Ikeda H., Ishikawa J., Hanamoto A., Shinose M., Kikuchi H., Shiba T.,
RA Sakaki Y., Hattori M., Omura S.;
RT "Complete genome sequence and comparative analysis of the industrial
RT microorganism Streptomyces avermitilis.";
RL Nat. Biotechnol. 21:526-531(2003).
CC -!- FUNCTION: Part of a binding-protein-dependent transport system.
CC Probably responsible for the translocation of the substrate across
CC the membrane (By similarity).
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (By similarity).
CC -!- SIMILARITY: Belongs to the binding-protein-dependent transport
CC system permease family.
DR EMBL; AP005029; BAC69805.1; -.

```

DR GO; GO:0016021; C: integral to membrane; IEA.
 DR GO; GO:0005351; F: sugar porter activity; IEA.
 DR GO; GO:0005215; F: transporter activity; IEA.
 DR GO; GO:0006810; P: transport; IEA.
 DR InterPro; IPR000515; BPD_transp.
 DR Pfam; PF00528; BPD_transp_1; 1.
 DR PROSITE; PS0928; ABC_TM1; 1.
 KW Complete proteome; Sugar transport; Transmembrane; Transport.
 SQ SEQUENCE 296 AA; 32715 MW; 70374AD2BB86FFB8A CRC64;

Query Match 88.9%; Score 48; DB 2; Length 296;
 Best Local Similarity 85.7%; Pred. No. 26;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYGGYW 7
 |||||
 Db 198 GYGAYW 204

RESULT 8
 Q8U043 PRELIMINARY; PRT; 363 AA.
 AC Q8U043;
 DT 01-JUN-2002 (TrEMBLrel. 21, Last Created)
 DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Iron (III) ABC transporter, ATP-binding protein.
 GN OrderedLocusNames=PF1774;
 OS Pyrococcus furiosus.
 OC Archaea; Euryarchaeota; Thermococci; Thermococcales; Thermococcaceae;
 OC Pyrococcus.
 OX NCBI_TaxID=2261;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=Vc1 / DSM 3638 / ATCC 43587 / JCM 8422;
 RA Weiss R.B., Dunn D.M., Robb F.T., Brown J.R.;
 RT "The complete sequence of the Pyrococcus furiosus genome."
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE010275; AAL81898.1; -.
 DR GO; GO:0005381; F: iron ion transporter activity; IEA.
 DR GO; GO:0006827; P: high affinity iron ion transport; IEA.
 DR Pfam; PF01497; Peripla_BP_2; 1.
 KW Complete proteome.
 SQ SEQUENCE 363 AA; 40352 MW; D36B5D4A5AE6CBC CRC64;

Query Match 87.0%; Score 47; DB 2; Length 363;
 Best Local Similarity 85.7%; Pred. No. 43;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYGGYW 7
 |||||
 Db 222 GYWDGYW 228

RESULT 9
 Q8A8U4 PRELIMINARY; PRT; 214 AA.
 AC Q8A8U4;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=BT1073;
 OS Bacteroides thetaiotaomicron.
 OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
 OC Bacteroidaceae; Bacteroides.
 OX NCBI_TaxID=818;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=VPI-5482 / ATCC 29148;
 RX MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
 RA Xu J., Bjursell M.K., Hinrod J., Deng S., Carmichael L.K.,
 Chiang H.C., Hooper L.V., Gordon J.I.;

RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis."
 RL Science 299:2074-2076(2003).
 DR EMBL; AE016930; AAO76180.1; -.
 KW Complete proteome.
 SQ SEQUENCE 214 AA; 24735 MW; 84ABB75A740D226B CRC64;

Query Match 85.2%; Score 46; DB 2; Length 214;
 Best Local Similarity 85.7%; Pred. No. 38;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYGGYW 7
 |||||
 Db 130 GYGGPYW 136

RESULT 10
 Q8YXK8 PRELIMINARY; PRT; 68 AA.
 AC Q8YXK8;
 DT 01-MAR-2002 (TrEMBLrel. 20, Created)
 DT 01-MAR-2002 (TrEMBLrel. 20, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Asr1204 protein.
 GN OrderedLocusNames=asr1204;
 OS Anabaena sp. (strain PCC 7120).
 OC Bacteria; Cyanobacteria; Nostocales; Nostocaceae; Nostoc.
 OX NCBI_TaxID=103690;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC MEDLINE=21595285; PubMed=11759840;
 RA Kaneko T., Nakamura Y., Wolk C.P., Kuritz T., Sasamoto S.,
 RA Watanabe A., Iriguchi M., Ishikawa A., Kawashima K., Kimura T.,
 RA Kishida Y., Kohara M., Matsumoto M., Mateuno A., Muraki A.,
 RA Nakazaki N., Shimpoto S., Sugimoto M., Takazawa M., Yamada M.,
 RA Yasuda M., Tabata S.;
 RT "Complete genomic sequence of the filamentous nitrogen-fixing
 cyanobacterium Anabaena sp. strain PCC 7120."
 RL DNA Res. 8:205-213(2001).
 DR EMBL; AF003585; BAB73161.1; -.
 DR FIR; A11956; A11956.
 KW Complete proteome.
 SQ SEQUENCE 68 AA; 7944 MW; BE31A1488358C962 CRC64;

Query Match 83.3%; Score 45; DB 2; Length 68;
 Best Local Similarity 85.7%; Pred. No. 20;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYGGYW 7
 |||||
 Db 61 GRWGGYW 67

RESULT 11
 Q9VZB1 PRELIMINARY; PRT; 106 AA.
 AC Q9VZB1;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE CG13721-PA.
 GN ORFNames=CG13721;
 OS Drosophila melanogaster (Fruit fly).
 OC Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 OC Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 OC Ephydroidea; Drosophilidae; Drosophila.
 OX NCBI_TaxID=7227;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=20196006; PubMed=10731132; DOI=10.1126/science.287.5461.2185;
 RA Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
 RA Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
 RA George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
 RA Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,

RA Brandon R.C., Rogers Y.H., Blazej R.G., Champe M., Pfeiffer B.D.,
 Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Gabor G.L.,
 RA April J.F., Agbayani A., An H.J., Andrews-Pfankuch C., Baldwin D.,
 RA Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
 RA Beeson K.V., Benos P.V., Berman B.P., Bhandari D., Bolehakov S.,
 RA Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
 RA Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
 RA Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
 RA de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
 RA Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
 RA Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
 RA Foeller C., Gabriellian A.E., Garg N.S., Gelbart W.M., Glasser K.,
 RA Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
 RA Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
 RA Hostin D., Houston K.A., Howland T.J., Wei M.H., Ibegwam C.,
 RA Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
 RA Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
 RA Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
 RA Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
 RA Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
 RA Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
 RA Nelson D.R., Nelson K.A., Nixon K., Nusser D.R., Pacleb J.M.,
 RA Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
 RA Reinert K., Remington K., Saunders R.D., Scheeler F., Shen H.,
 RA Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
 RA Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
 RA Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
 RA Wang Z.Y., Wasserman D.A., Weinstock G.M., Weissbach J.,
 RA Williams S.M., Woodgett W., Worley K.C., Wu D., Yang S., Yao Q.A., Ye J.,
 RA Yeh R.F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
 RA Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Smith H.O.,
 RA Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
 RA "The genome sequence of *Drosophila melanogaster*.";
 RA Science 287:2185-2195(2000).
 RN [2]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=22426065; PubMed=12537568;
 RA Celniker S.E., Wheeler D.A., Kronmiller B., Carlson J.W., Halpern A.,
 RA Patel S., Adams M., Champe M., Dugan S.P., Frise E., Hodgson A.,
 RA George R.A., Hoskins R.A., Lavery T., Muzny D.M., Nelson C.R.,
 RA Pacleb J.M., Park S., Pfeiffer B.D., Richards S., Sodergren E.J.,
 RA Svirskas R., Taber P.E., Wan K., Stapleton M., Sutton G.G., Venter C.,
 RA Weinstock G., Scherer S.E., Myers E.W., Gibbs R.A., Rubin G.M.;
 RA "Finishing a whole-genome shotgun: Release 3 of the *Drosophila*
 RA melanogaster euchromatic genome sequence.";
 RA Genome Biol. 3:RESEARCH0079-RESEARCH0079(2002).
 RN [3]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=22426070; PubMed=12537573;
 RA Kaminker J.S., Bergman C.M., Kronmiller B., Carlson J., Svirskas R.,
 RA Patel S., Frise E., Wheeler D.A., Lewis S.E., Rubin G.M.,
 RA Ashburner M., Celniker S.E.;
 RA "The transposable elements of the *Drosophila melanogaster* euchromatin:
 RA a genomics perspective.";
 RA Genome Biol. 3:RESEARCH0084-RESEARCH0084(2002).
 RN [4]
 RN SEQUENCE FROM N.A.
 RX MEDLINE=22426069; PubMed=12537572;
 RA Misra S., Crosby M.A., Mungall C.J., Matthews B.B., Campbell K.S.,
 RA Hradecky P., Huang Y., Kaminker J.S., Millburn G.H., Prochnik S.E.,
 RA Smith C.D., Tupy J.L., Whitfield E.J., Bayraktaroglu L., Berman B.P.,
 RA Battencourt B.R., Celniker S.E., de Grey A.D., Drysdale R.A.,
 RA Harris N.L., Richter J., Russo S., Schroeder A.J., Shu S.Q.,
 RA Stapleton M., Yamada C., Ashburner M., Gelbart W.M., Rubin G.M.,
 RA Lewis S.E.;
 RA "Annotation of the *Drosophila melanogaster* euchromatic genome: a
 RA systematic review.";
 RA Genome Biol. 3:RESEARCH0083-RESEARCH0083(2002).
 RN [5]
 RN SEQUENCE FROM N.A.
 RG FlyBase;
 RN Submitted (SEP-2002) to the EMBL/GenBank/DBJ databases.
 RN [6]

RP SEQUENCE FROM N.A.
 RG FlyBase;
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AE003481; AAF47913.1; -;
 DR IntAct; Q9VZB1; -;
 DR FlyBase; FBgn0035554; CG13721.
 SQ SEQUENCE 106 AA; 11918 MW; BA74FF7C320474EE CRC64;
 Query Match 83.3%; Score 45; DB 2; Length 106;
 Best Local Similarity 85.7%; Pred. No. 29;
 Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;
 Qy 1 GYWGYYW 7
 Db 68 GYWGYYW 74
 ||:||||
 ||:||||
 RESULT 12
 CDSA_PSEAE STANDARD; PRT; 271 AA.
 ID CDSA_PSEAE
 AC Q59640;
 DT 01-NOV-1997 (Rel. 35, Created)
 DT 15-OCT-2001 (Rel. 40, Last sequence update)
 DT 05-JUL-2004 (Rel. 44, Last annotation update)
 DE Phosphatidate cytidyltransferase (EC 2.7.7.41) (CDP-diglyceride
 DE synthetase) (CDP-diglyceride pyrophosphorylase) (CDP-diacylglycerol
 DE synthase) (CDS) (CTP:phosphatidate cytidyltransferase) (CDP-DAG
 DE synthase) (CDP-DG synthetase).
 GN NamescdsA; Synonyms=cds; OrderedLocusNames=PA3651;
 GN Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 ON NCBI_TaxID=287;
 OX [1]
 RN SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=96257274; PubMed=8654980; DOI=10.1016/0378-1119(96)00009-1;
 RA Taguchi K., Fukutomi H., Kuroda A., Kato J., Ohtake H.;
 RT "Cloning of the *Pseudomonas aeruginosa* gene encoding CDP-diglyceride
 RT synthetase.";
 RL Gene 172:165-166(1996).
 RN [2]
 RN SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goitry L., Tolentino E., Westbrook-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Saier M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of *Pseudomonas aeruginosa* PA01, an
 RT opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 CC -|- CATALYTIC ACTIVITY: CTP + phosphatidate = diphosphate + CDP-
 CC diacylglycerol.
 CC -|- PATHWAY: Phospholipid biosynthesis.
 CC -|- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
 CC -|- SIMILARITY: Belongs to the CDS family.
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 DR EMBL; D50811; BAA09437.1; -;
 DR EMBL; AB004785; AAG07039.1; -;
 DR FIC; F83188; F83188.
 DR FIC; JC4832; JC4832.
 DR InterPro; IPR000374; PC_trans.

DR Pfam; PF01148; CTP transf 1; 1.
 DR PROSITE; PS01315; CDS; 1.
 KW Complete proteome; Inner membrane; Nucleotidyltransferase;
 KW Phospholipid biosynthesis; Transferase; Transmembrane.
 FT TRANSMEM 12 32 Potential.
 FT TRANSMEM 53 73 Potential.
 FT TRANSMEM 75 95 Potential.
 FT TRANSMEM 111 131 Potential.
 FT TRANSMEM 136 156 Potential.
 FT TRANSMEM 174 194 Potential.
 FT TRANSMEM 199 219 Potential.
 FT TRANSMEM 251 271 Potential.
 FT CONFLICT 131 133 WPL -> VAA (in Ref. 1).
 SQ SEQUENCE 271 AA; 28856 MW; 5025059C3F1A64C7 CRC64;

Query Match 83.3%; Score 45; DB 1; Length 271;
 Best Local Similarity 85.7%; Pred. No. 65;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYGGYV 7
 |||||
 Db 102 GYGGRW 108

RESULT 13

Q7SD72 PRELIMINARY; PRT; 762 AA.
 AC Q7SD72;
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Predicted protein.
 GN Name=NCU02790.1;
 OS Neurospora crassa.
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
 OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
 OX NCBI_TaxID=5141;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=OR74A;
 RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
 RA Jaffe D., FitzHugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
 RA Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
 RA Qui D., Ianakiev P., Pedersen D., Nelson M., Washburne M.,
 RA Selitrenkoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,
 RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
 RA Roy A., Foley K., Naylor J., Thomann N., Barrett R., Gnerre S.,
 RA Kamal M., Kanvasselis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
 RA Krystofova S., Rasmussen C., Metzberg R.L., Perkins D.D., Kroken S.,
 RA Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmani S.A.,
 RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
 RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
 RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbole D.J., Freitag M.,
 RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.;
 RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa."
 RL Nature 0:0-0 (2003).
 CC -1- CAUTION: The sequence shown here is derived from an
 CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
 CC preliminary data.
 DR EMBL; AABX0100062; EAA34710.1; -;
 DR InterPro; IPR011014; Mcds_transmembr.
 SQ SEQUENCE 762 AA; 84704 MW; 2C7DD1515E1902FA CRC64;

Query Match 83.3%; Score 45; DB 2; Length 762;
 Best Local Similarity 71.4%; Pred. No. 1.6e+02;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYGGYV 7
 |||||
 Db 554 GYGGRW 560

RESULT 14

Query Match 77.8%; Score 42; DB 2; Length 210;
 Best Local Similarity 71.4%; Pred. No. 1.4e+02;

Q6LSS8

ID Q6LSS8 PRELIMINARY; PRT; 302 AA.
 AC Q6LSS8;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=PBPA1237;
 OS Photobacterium profundum (Photobacterium sp. (strain SS9)).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
 OC Vibrionaceae; Photobacterium.
 OX NCBI_TaxID=74109;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Vezzi A., Campanaro S., D'Angelo M., Simonato F., Vitulo N., Lauro F.,
 RA Cestari A., Malacrida G., Simionati B., Cannata N., Bartlett D.,
 RA Valle G.;
 RT "Genome analysis of Photobacterium profundum reveals the complexity of
 RT high pressure adaptations.";
 RL Submitted (MAR-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; CR378667; CAG19648.1; -;
 KW Complete proteome.
 SQ SEQUENCE 302 AA; 34200 MW; 708FA80ED34BADE7 CRC64;

Query Match 81.5%; Score 44; DB 2; Length 302;
 Best Local Similarity 83.3%; Pred. No. 99;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

Qy 2 YWGGYV 7
 |||||
 Db 199 YWGGFW 204

RESULT 15

Q88RD4 PRELIMINARY; PRT; 210 AA.
 AC Q88RD4;
 DT 01-JUN-2003 (TrEMBLrel. 24, Created)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
 DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
 DE Transporter, lyse family.
 GN OrderedLocusNames=PP0198;
 OS Pseudomonas putida (strain KT2440).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=160486;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX MEDLINE=22423060; PubMed=12534463;
 RA Nelson K.E., Weinel C., Paulsen I.T., Dodson R.J., Hilbert H.,
 RA Martins dos Santos V.A.P., Fouts D.E., Gill S.R., Pop M., Holmes M.,
 RA Brinkac L.M., Beanan M.J., DeBoy R.T., Daugherty S.C., Kolonay J.F.,
 RA Madupu R., Nelson W.C., White O., Peterson J.D., Khouri H.M.,
 RA Hance I., Chris Lee P., Holtzapple E.K., Scanlan D., Tran K.,
 RA Moazzaz A., Uterback T.R., Rizzo M., Lee K., Kosack D., Moestl D.,
 RA Wedler H., Lauber J., Stjepandic D., Hoheisel J., Straetz M., Heim S.,
 RA Kiewitz C., Eissen J.A., Timmis K.N., Duesterhoeft A., Tummliar B.,
 RA Fraser C.M.;
 RT "Complete genome sequence and comparative analysis of the
 RT metabolically versatile Pseudomonas putida KT2440.";
 RL Environ. Microbiol. 4:799-808 (2002).
 DR EMBL; AE016774; AAN65831.1; -;
 DR TIGR; PP0198; -;
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0005293; F:lysine permease activity; IEA.
 DR GO; GO:0006865; P:amino acid transport; IEA.
 DR InterPro; IPR001123; LysE.
 DR Pfam; PF01810; LysE; 1.
 KW Complete proteome.
 SQ SEQUENCE 210 AA; 22827 MW; 89E800A87428770A CRC64;

Query Match 77.8%; Score 42; DB 2; Length 210;
 Best Local Similarity 71.4%; Pred. No. 1.4e+02;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYGGYW 7
|:|
Db 35 GFWRGYW 41

Search completed: April 18, 2005, 19:58:23
Job time : 65.3488 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 19:18:17 ; Search time 71.7907 Seconds
(without alignments)
37.711 Million cell updates/sec

Title: SEQ1

Perfect score: 54

Sequence: 1 gywgyw 7

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : A_Geneseq_16Dec04:.*
1: Geneseqp1980s:.*
2: Geneseqp1990s:.*
3: Geneseqp2000s:.*
4: Geneseqp2001s:.*
5: Geneseqp2002s:.*
6: Geneseqp2003as:.*
7: Geneseqp2003bs:.*
8: Geneseqp2004s:.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	54	100.0	339	7	ABO78095
2	54	100.0	440	6	ABU20398
3	45	83.3	106	4	ABB68525
4	45	83.3	500	7	ABO72468
5	43.5	80.6	8	2	AAU03714
6	43.5	80.6	8	2	AAU03716
7	43.5	80.6	8	7	ADG94002
8	43.5	80.6	8	8	ADL38013
9	43.5	80.6	9	3	AAU76816
10	43.5	80.6	131	4	AAU11208
11	42	77.8	118	8	ADN49403
12	42	77.8	211	6	AAU39959
13	42	77.8	243	2	AAU09436
14	42	77.8	447	6	ABM67619
15	41	75.9	20	5	AAU90607
16	41	75.9	20	6	ADA04444
17	41	75.9	20	7	ADH95657
18	41	75.9	20	8	ADL68348
19	41	75.9	20	8	ADM38193
20	41	75.9	26	6	ADA04885
21	41	75.9	26	7	ADH96098
22	41	75.9	26	8	ADL68789
23	41	75.9	26	8	ADM38634
24	41	75.9	39	5	AAU90645
25	41	75.9	40	6	ADA04482

26	41	75.9	40	7	ADH95695	Adh95695 Insulin g
27	41	75.9	40	8	ADL68386	Adl68386 IGF-IR/IR
28	41	75.9	40	8	ADM38231	Adm38231 Anti-IGF-
29	41	75.9	105	4	ABB50792	Abb50792 Human sec
30	41	75.9	105	6	ABO45049	AbO45049 Novel hum
31	41	75.9	105	7	ABO26529	AbO26529 Protein a
32	41	75.9	290	6	ABP77676	Abp77676 N. gonorr
33	41	75.9	396	6	ABU19932	Abu19932 Protein e
34	41	75.9	396	6	ABU22337	Abu22337 Protein e
35	41	75.9	400	8	ADN18476	Adn18476 Bacterial
36	41	75.9	403	6	ABU21628	Abu21628 Protein e
37	40	74.1	464	3	AAG16993	Aag16993 Arabidops
38	40	74.1	543	7	ADE49759	Ade49759 Red rice
39	40	74.1	552	7	ADE49762	Ade49762 Rice incr
40	40	74.1	552	7	ADE49775	Ade49775 Rice incr
41	40	74.1	552	7	ADE49778	Ade49778 Rice incr
42	40	74.1	552	7	ADE49765	Ade49765 Rice incr
43	40	74.1	552	7	ADE49768	Ade49768 Rice incr
44	40	74.1	552	7	ADE49755	Ade49755 Red rice
45	40	74.1	552	7	ADE49771	Ade49771 Rice incr

ALIGNMENTS

RESULT 1
ABO78095
ID ABO78095 standard; protein; 339 AA.
XX ABO78095;
XX
XX
XX 29-JUL-2004 (first entry)
XX Pseudomonas aeruginosa polypeptide #10270.
XX
XX Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX Pseudomonas aeruginosa.
XX
XX US6551795-B1.
XX
XX 22-APR-2003.
XX
XX 18-FEB-1999; 99US-00252991.
XX
XX 18-FEB-1998; 98US-0074788P.
XX 27-JUL-1998; 98US-0094190P.
XX
XX (GENO-) GENOME THERAPEUTICS CORP.
XX
XX Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
XX WPI; 2003-615309/58.
XX N-PSDB; ABD11666.
XX
XX Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
XX useful as molecular targets for diagnostics, prophylaxis and treatment of
XX pathological conditions resulting from bacterial infection.
XX
XX Disclosure; SEQ ID NO 26941; 455pp; English.

Rubenfield MJ, Nolling J, Deloughery C, Bush D;

WPI; 2003-615309/58.
N-PSDB; ABD11666.

Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
useful as molecular targets for diagnostics, prophylaxis and treatment of
pathological conditions resulting from bacterial infection.

Disclosure; SEQ ID NO 26941; 455pp; English.

The invention relates to Pseudomonas aeruginosa polypeptides and the
polynucleotides encoding them. The sequences are useful in diagnosis and
therapy of pathological conditions, as molecular targets for diagnostics,
prophylaxis and treatment of pathological conditions resulting from a
bacterial infection, for evaluating a compound, such as a polypeptide,
for the ability to bind a P. aeruginosa nucleic acid, as components of
effective antibacterial targets, as targets for antibacterial drugs,
including anti-P. aeruginosa drugs, as templates for recombinant
production of P. aeruginosa-derived peptides or polypeptides, as target
components for diagnosis and/or treatment of P. aeruginosa-caused
infection, and in detection of P. aeruginosa sequences or other sequences
of Pseudomonas species using biochip technology. Sequences ABO67826-

CC AB084396 represent P. aeruginosa polypeptides of the invention. Note: The
 CC sequence data for this patent did not form part of the printed
 CC specification but was obtained in electronic format from USPTO at
 CC seqdata.uspto.gov/sequence.html
 XX
 SQ Sequence 339 AA;

Query Match 100.0%; Score 54; DB 7; Length 339;
 Best Local Similarity 100.0%; Pred. No. 5.5;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYGGYV 7
 |||||
 Db 261 GYGGYV 267

RESULT 2
 ABU20398
 ID ABU20398 standard; protein; 440 AA.
 XX
 AC ABU20398;
 XX
 DT 19-JUN-2003 (first entry)
 XX
 DE Protein encoded by Prokaryotic essential gene #5925.
 XX
 KW Antisense; prokaryotic essential gene; cell proliferation; drug design.
 XX
 OS Bacteroides fragilis.
 XX
 PN WO200277183-A2.
 XX
 PD 03-OCT-2002.
 XX
 PF 21-MAR-2002; 2002WO-US009107.
 XX
 PR 21-MAR-2001; 2001US-00815242.
 PR 06-SEP-2001; 2001US-00948993.
 PR 28-OCT-2001; 2001US-0342923P.
 PR 08-FEB-2002; 2002US-00072851.
 PR 06-MAR-2002; 2002US-0362699P.
 XX
 PA (ELIT-) ELITRA PHARM INC.
 XX
 PI Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
 PI Wali D, Trawick JD, Carr GU, Yamamoto R, Forsyth RA, Xu HH;
 XX
 DR WPI; 2003-029926/02.
 DR N-PSDB; ACA24268.
 XX
 PT New antisense nucleic acids, useful for identifying proteins or screening
 PT for homologous nucleic acids required for cellular proliferation to
 PT isolate candidate molecules for rational drug discovery programs.
 XX
 PS Claim 25; SEQ ID NO 48322; 1766pp; English.
 XX

The invention relates to an isolated nucleic acid comprising any one of
 the 6213 antisense sequences given in the specification where expression
 of the nucleic acid inhibits proliferation of a cell. Also included are:
 (1) a vector comprising a promoter operably linked to the nucleic acid
 encoding a polypeptide whose expression is inhibited by the antisense
 nucleic acid; (2) a host cell containing the vector; (3) an isolated
 polypeptide or its fragment whose expression is inhibited by the
 antisense nucleic acid; (4) an antibody capable of specifically binding
 the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
 proliferation or the activity of a gene in an operon required for
 proliferation; (7) identifying a compound that influences the activity of
 the gene product or that has an activity against a biological pathway
 required for proliferation, or that inhibits cellular proliferation; (8)
 identifying a gene required for cellular proliferation or the biological
 pathway in which a proliferation-required gene or its gene product lies
 or a gene on which the test compound that inhibits proliferation of an
 organism acts; (9) manufacturing an antibiotic; (10) profiling a

CC compound's activity; (11) a culture comprising strains in which the gene
 CC product is overexpressed or underexpressed; (12) determining the extent
 CC to which each of the strains is present in a culture or collection of
 CC strains; or (13) identifying the target of a compound that inhibits the
 CC proliferation of an organism. The antisense nucleic acids are useful for
 CC identifying proteins or screening for homologous nucleic acids required
 CC for cellular proliferation to isolate candidate molecules for rational
 CC drug discovery programs, or for screening homologous nucleic acids
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
 CC the target prokaryotic essential genes. Note: the sequence data for this
 CC patent did not form part of the printed specification, but was obtained
 CC in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences
 XX
 SQ Sequence 440 AA;

Query Match 100.0%; Score 54; DB 6; Length 440;
 Best Local Similarity 100.0%; Pred. No. 6.9;
 Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GYGGYV 7
 |||||
 Db 212 GYGGYV 218

RESULT 3
 ABB68525
 ID ABB68525 standard; protein; 106 AA.
 XX
 AC ABB68525;
 XX
 DT 26-MAR-2002 (first entry)
 XX
 DE Drosophila melanogaster polypeptide SEQ ID NO 32367.
 XX
 KW Drosophila; developmental biology; cell signalling; insecticide;
 KW pharmaceutical.
 XX
 OS Drosophila melanogaster.
 XX
 PN WO200171042-A2.
 XX
 PD 27-SEP-2001.
 XX
 PF 23-MAR-2001; 2001WO-US009231.
 XX
 PR 23-MAR-2000; 2000US-0191637P.
 PR 11-JUL-2000; 2000US-00614150.
 XX
 PA (PEKE) PE CORP NY.
 XX
 PI Venter JC, Adams M, Li PWD, Myers EW;
 XX
 DR WPI; 2001-656860/75.
 DR N-PSDB; ABL12628.
 XX
 PT New isolated nucleic acid detection reagent for detecting 1000 or more
 PT genes from *Drosophila* and for elucidating cell signalling and cell-cell
 PT interactions.
 XX
 PS Disclosure; SEQ ID NO 32367; 21pp + Sequence Listing; English.
 XX

The invention relates to an isolated nucleic acid detection reagent
 capable of detecting 1000 or more genes from *Drosophila*. The invention is
 useful in developmental biology and in elucidating cell signalling and
 cell-cell interactions in higher eukaryotes for the development of
 insecticides, therapeutics and pharmaceutical drugs. The invention
 discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA
 sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-
 ABB72072). The sequence data for this patent did not form part of the
 printed specification, but was obtained in electronic format directly
 from WIPO at ftp.wipo.int/pub/published_pct_sequences

```

XX SQ Sequence 106 AA;
Query Match      83.3%; Score 45; DB 4; Length 106;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYGGYW 7
   ||:||||
Db 68 GYGGYW 74

RESULT 4
ID ABO72468 standard; protein; 500 AA.
AC ABO72468;
XX
DT 29-JUL-2004 (first entry)
DE Pseudomonas aeruginosa polypeptide #4643.
KW Bacterial infection; Pseudomonas aeruginosa infection; antibacterial.
XX
OS Pseudomonas aeruginosa.
XX
PN US6551795-B1.
XX
PD 22-APR-2003.
XX
PF 18-FEB-1999; 99US-00252991.
XX
PR 18-FEB-1998; 98US-0074788P.
PR 27-JUL-1998; 98US-0094190P.
XX
PA (GENO-) GENOME THERAPEUTICS CORP.
XX
PI Rubenfield MJ, Nolling J, Deloughery C, Bush D;
XX
DR WPI; 2003-615309/58.
DR N-PSDB; ABD06039.
XX
PT Novel isolated nucleic acid encoding Pseudomonas aeruginosa polypeptide,
PT useful as molecular targets for diagnostics, prophylaxis and treatment of
PT pathological conditions resulting from bacterial infection.
XX
PS Disclosure; SEQ ID NO 2124; 455pp; English.
XX
CC The invention relates to Pseudomonas aeruginosa polypeptides and the
CC polynucleotides encoding them. The sequences are useful in diagnosis and
CC therapy of pathological conditions, as molecular targets for diagnostics,
CC prophylaxis and treatment of pathological conditions resulting from a
CC bacterial infection, for evaluating a compound, such as a polypeptide,
CC for the ability to bind a P. aeruginosa nucleic acid, as components of
CC effective antibacterial targets, as targets for antibacterial drugs,
CC including anti-P. aeruginosa drugs, as templates for recombinant
CC production of P. aeruginosa-derived peptides or polypeptides, as target
CC components for diagnosis and/or treatment of P. aeruginosa-caused
CC infection, and in detection of P. aeruginosa sequences or other sequences
CC of Pseudomonas species using biochip technology. Sequences ABO67826-
CC ABO84396 represent P. aeruginosa polypeptides of the invention. Note: The
CC sequence data for this patent did not form part of the printed
CC specification but was obtained in electronic format from USPTO at
CC seqdata.uspto.gov/sequence.html
XX
SQ Sequence 500 AA;
Query Match      83.3%; Score 45; DB 7; Length 500;
Best Local Similarity 85.7%; Pred. No. 1.5e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYGGYW 7
   |||||

```

```

Db 331 GYGGWRW 337

RESULT 5
AAY03714
ID AAY03714 standard; peptide; 8 AA.
XX
AC AAY03714;
XX
DT 08-JUN-1999 (first entry)
XX
DE Fluorine-18 (F-18) labeled peptide 1.
XX
KW 18F radionuclide; targeting vector; positron emission tomography; F-18;
KW radiolabeling; thiol; fluorine-18.
XX
OS Synthetic.
XX
XX Key Location/Qualifiers
FT Misc-difference 1 /note= "optionally has a free or protected amino acid
FT group"
FT Misc-difference 2 /note= "D-form residue"
FT Misc-difference 3 /note= "D-form residue"
FT Misc-difference 5 /note= "D-form residue"
FT /note= "D-form residue; optionally has a free or
FT protected amino acid group"
FT Misc-difference 7 /note= "D-form residue"
FT Misc-difference 8 /note= "D-form residue"
XX
PN WO9911590-A1.
XX
PD 11-MAR-1999.
XX
PF 03-SEP-1998; 98WO-US018268.
XX
PR 03-SEP-1997; 97US-0057485P.
XX
PA (IMMU-) IMMUNOMEDICS INC.
XX
PI Griffiths GL;
XX
DR WPI; 1999-228967/19.
XX
PT Radiolabeling thiol-containing peptides with fluorine-18.
XX
PS Claim 13; Page 15; 22pp; English.
XX
CC The invention relates to a method for incorporating 18F radionuclide into
CC peptide-containing targeting vectors for use in clinical positron
CC emission tomography. Radiolabeling thiol-containing peptides with
CC fluorine-18 (F-18) comprises reacting a peptide comprising a free thiol
CC group with a labeling reagent of formula: 18F-(CH2)m-CR1R2-(CH2)n-X, or a
CC fluorinated alkene in which at least one of the two double bonded carbon
CC atoms bears at least one leaving group comprising I, Br, Cl, azide, or
CC tosylate, mesylate, nosylate or triflate. n, m = 0-2; n+m = 0-2; X = I,
CC Br, Cl, azide, tosylate, mesylate, nosylate, triflate, maleimide
CC (optionally substituted by 1-2 alkyl) or 3-sulfonamide; R1, R2 = I,
CC Br, Cl, azide, tosylate, mesylate, nosylate, triflate, H, CONH2, COOH,
CC OH, sulfonic acid, tertiary amine, quaternary ammonium, alkyl (optionally
CC substituted by CONH2, COOH, OH, sulfonic acid, tertiary amine or
CC quaternary ammonium), COOR', CONR'2 or COR'; and R' = 1-6C alkyl or
CC phenyl. The method is used for Radiolabeling peptide-containing targeting
CC vectors such as proteins, antibodies, antibody fragments and receptor-
CC targeted peptides for use in routine clinical positron emission
CC tomography. The method is simple and efficient. The method uses the
CC unique property of the free thiol groups which are rapidly alkylated at
CC neutral pH and moderate temperature. Sequences AAY03714-716 represent
CC examples of F-18 labeled peptides used in the method of detecting a

```

CC tissue
 XX Sequence 8 AA;
 SQ

Query Match 80.6%; Score 43.5; DB 2; Length 8;
 Best Local Similarity 87.5%; Pred. No. 1.8e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 GYWG-GYW 7
 |||||
 Db 1 GYWGKGW 8

RESULT 6
 AAY03716
 ID AAY03716 standard; peptide; 8 AA.
 XX
 AC AAY03716;
 XX
 DT 08-JUN-1999 (first entry)
 XX
 DE Fluorine-18 (F-18) labeled peptide 3.
 XX
 KW 18F radionuclide; targeting vector; positron emission tomography; F-18;
 KW radiolabeling; thiol; fluorine-18.
 XX
 OS Synthetic.
 XX

Key Location/Qualifiers
 FT Modified-site 1 /note= "N-terminal acetylation"
 FT
 FT Misc-difference 2 /note= "D-form residue; D-iodo-Tyr"
 FT
 FT Misc-difference 3 /note= "D-form residue"
 FT
 FT Misc-difference 5 /note= "D-form residue; optionally acetylated"
 FT
 FT Misc-difference 7 /note= "D-form residue; D-iodo-Tyr"
 FT
 FT Misc-difference 8 /note= "D-form residue"
 FT
 XX WO9911590-A1.
 PN
 XX
 PD 11-MAR-1999.
 XX
 PF 03-SEP-1998; 98WO-US018268.
 XX
 PR 03-SEP-1997; 97US-0057485P.
 XX
 XX (IMMU-) IMMUNOMEDICS INC.
 PA
 XX Griffiths GL;
 PI
 XX WPI; 1999-228967/19.
 DR
 XX Radiolabeling thiol-containing peptides with fluorine-18.
 XX
 PS Claim 15; Page 15; 22pp; English.
 XX

The invention relates to a method for incorporating 18F radionuclide into peptide-containing targeting vectors for use in clinical positron emission tomography. Radiolabeling thiol-containing peptides with fluorine-18 (F-18) comprises reacting a peptide comprising a free thiol group with a labeling reagent of formula: 18F-(CH₂)_m-CH(R₂)-(CH₂)_n-X, or a fluorinated alkene in which at least one of the two double bonded carbon atoms bears at least one leaving group comprising I, Br, Cl, azide, tosylate, mesylate, nosylate or triflate. n, m = 0-2; n+m = 0-2; X = I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, maleimide (optionally substituted by 1-2 alkyl) or 3-sulfomaleimide; R₁, R₂ = I, Br, Cl, azide, tosylate, mesylate, nosylate, triflate, H, CONH₂, COOH, OH, sulfonic acid, tertiary amine, quaternary ammonium, alkyl (optionally substituted by CONH₂, COOH, OH, sulfonic acid, tertiary amine or

CC quaternary ammonium), COOR', CONR'2 or COR'; and R' = 1-6C alkyl or phenyl. The method is used for Radiolabeling peptide-containing targeting vectors such as proteins, antibodies, antibody fragments and receptor-targeted peptides for use in routine clinical positron emission tomography. The method is simple and efficient. The method uses the unique property of the free thiol groups which are rapidly alkylated at neutral pH and moderate temperature. Sequences AAY03714-716 represent examples of F-18 labeled peptides used in the method of detecting a tissue
 XX
 XX Sequence 8 AA;
 SQ

Query Match 80.6%; Score 43.5; DB 2; Length 8;
 Best Local Similarity 87.5%; Pred. No. 1.8e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
 QY 1 GYWG-GYW 7
 |||||
 Db 1 GYWGKGW 8

RESULT 7
 ADG94002
 ID ADG94002 standard; peptide; 8 AA.
 XX
 AC ADG94002;
 XX
 DT 11-MAR-2004 (first entry)
 XX
 DE Iodinated immunogenic peptide.
 XX
 KW Iodinated immunogenic peptide; multi-specific antibody;
 KW polymer conjugate; tumour; cytostatic; photodynamic therapy.
 XX
 OS Synthetic.
 XX

Key Location/Qualifiers
 FT Modified-site 1 /note= "Acetylated"
 FT
 FT Misc-difference 2 /note= "D-form residue and optionally iodinated"
 FT
 FT Misc-difference 3 /note= "D-form residue"
 FT
 FT Misc-difference 5 /note= "Acetylated D-form residue"
 FT
 FT Misc-difference 7 /note= "D-form residue and optionally iodinated"
 FT
 FT Misc-difference 8 /note= "D-form residue"
 FT
 XX US2003026764-A1.
 PN
 XX
 PD 06-FEB-2003.
 XX
 PF 31-JUL-2002; 2002US-00209592.
 XX
 PR 31-JUL-2001; 2001US-0308605P.
 XX
 XX (IMMU-) IMMUNOMEDICS INC.
 PA
 XX Griffiths GL;
 PI
 XX WPI; 2003-801085/75.
 DR
 XX Targeting an agent towards a target site in a tissue, by administering in a tissue, a multi-specific antibody or its fragment and a polymer conjugate that binds to the capture arm of the multi-specific antibody.
 PT
 XX Disclosure; Page 8; 19pp; English.
 PS
 XX The invention relates to targeting an agent towards a target site in a tissue, comprising administering to the tissue, a multi-specific antibody or its fragment, comprising a targeting arm that binds to an antigen of

CC the target site and a capture arm that binds to a polymer conjugate, and
 CC administering a polymer conjugate that binds to the capture arm, the
 CC conjugate has a polymer conjugated to the agent such as therapeutic
 CC agent, a peptide, an enzyme and a labelled ligand. Also included is a kit
 CC useful for targeting a target site within a tissue in a subject or tissue
 CC sample comprising the above mentioned multi-specific antibody or its
 CC fragment and a polymer conjugate. The method is used for targeting an
 CC agent towards a target site in a tissue (e.g. a tumour). The method is
 CC also useful for therapeutic or diagnostic purposes and further in
 CC photodynamic therapy. The present sequence is an Iodinated immunogenic
 CC peptide used to raise antibodies which recognise only the iodinated
 CC peptide, used in the method of the invention.

XX Sequence 8 AA;

Query Match 80.6%; Score 43.5; DB 7; Length 8;
 Best Local Similarity 87.5%; Pred. No. 1.8e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

Qy 1 GYWG-GYW 7
 |||||
 Db 1 GYWGKYW 8

RESULT 8

ADL98013
 ID ADL98013 standard; peptide; 8 AA.

AC ADL98013;

DT 20-MAY-2004 (first entry)

DE Peptide haptan #1.

XX photodynamic diagnosis; cancer; tumour; cardiovascular lesion;

KW inflammatory disease; neurodegenerative disease; metabolic disease;

KW infectious disease; B-cell malignancy; Alzheimer's disease; amyloidosis;

KW autoimmune disease; bacterial infection; fungal infection;

KW parasitic infection; viral infection;

KW carcinoembryonic antigen-expressing tumour.

XX Synthetic.

XX Key Location/Qualifiers

FH Modified-site 1

FT /note= "N-terminal acetyl"

FT Modified-site 2 /note= "D form residue; optionally iodo Tyr"

FT Misc-difference 3

FT /note= "D form residue"

FT Modified-site 5

FT /note= "D form residue; acetylated"

FT Modified-site 7

FT /note= "D form residue; optionally iodo Tyr"

FT Misc-difference 8

FT /note= "D form residue"

FT US2004043030-A1.

PN 04-MAR-2004.

XX 09-JUN-2003; 2003US-00456580.

XX 31-JUL-2001; 2001US-0308605P.

XX 31-JUL-2002; 2002US-00209592.

XX (IMMU-) IMMUNOMEDICS INC.

XX Griffiths GL, Goldenberg DM, Hansen HJ;

PI WPI; 2004-313739/29.

XX Treating cancer and metabolic diseases by administering a multi-specific

PT antibody having a targeting arm that binds to an antigen and a capture
 PT arm that binds to a polymer conjugate comprising a therapeutic agent.
 XX Disclosure; Page 16; 24pp; English.

XX The invention relates to a method of diagnosing or treating a disease or
 disorder. The method involves administering to a tissue a multi-specific
 antibody (I) or antibody fragment, comprising a targeting arm that binds
 to an antigen on the target site, and a capture arm that binds to a
 polymer conjugate, and administering to the tissue a polymer conjugate
 that binds to the capture arm, the polymer conjugate comprising a polymer
 conjugated to a diagnostic or therapeutic agent. Also included is a
 method for photodynamic diagnosis or treatment of a disease or disorder;
 or intravascular or endoscopic method for diagnosing or treating a
 disease or disorder. The method is useful for diagnosing or treating a
 disease or disorder chosen from cancer (oesophageal, gastric, colonic,
 rectal, pancreatic, lung, breast, ovarian, urinary bladder, endometrial,
 cervical, testicular, renal, adrenal and liver cancer, solid tumour, B-
 cell malignancy or T-cell malignancy); cardiovascular lesion; an
 inflammatory disease; neurodegenerative disease; metabolic disease; and
 an infectious disease. The B-cell malignancy is chosen from indolent
 forms of B-cell lymphomas, aggressive forms of B-cell lymphomas, chronic
 lymphatic leukaemias, acute lymphatic leukaemias, and multiple myeloma.
 The solid tumour is chosen melanoma, carcinoma (preferably renal
 carcinoma, lung carcinoma, intestinal carcinoma, and stomach carcinoma),
 glioma and sarcoma. The cardiovascular lesion is chosen from infarct,
 clot, embolus, atherosclerotic plaque and ischaemia. The
 neurodegenerative disease is Alzheimer's disease. The metabolic disease
 is amyloidosis, where the antibody binds amyloid. The disease or disorder
 is displaced or ectopic normal tissue chosen from endometrium, thymus,
 spleen and parathyroid. The method can be used for normal tissue
 ablation, where the tissue is chosen from bone marrow and spleen. The
 disease or disorder is an autoimmune disease such as myasthenia gravis,
 lupus nephritis, lupus erythematosus, and rheumatoid arthritis, Class III
 autoimmune diseases such as immune-mediated thrombocytopenias, such as
 acute idiopathic thrombocytopenic purpura and chronic idiopathic
 thrombocytopenic purpura, dermatomyositis, Sjogren's syndrome, multiple
 sclerosis, Sydenham's chorea, myasthenia gravis, systemic lupus
 erythematosus, lupus nephritis, rheumatic fever, polyglandular syndromes,
 bullous pemphigoid, diabetes mellitus, Henoch-Schonlein purpura, post-
 streptococcal nephritis, erythema nodosum, Takayasu's arteritis, colitis,
 Addison's disease, rheumatoid arthritis, sarcoidosis, ulcerative colitis,
 erythema multiforme, IGA nephropathy, polyarteritis nodosa, ankylosing
 spondylitis, Goodpasture's syndrome, thromboangitis obliterans, primary
 biliary cirrhosis, Hashimoto's thyroiditis, thyrotoxicosis, scleroderma,
 chronic active hepatitis, polymyositis/dermatomyositis, polychondritis,
 pemphigus vulgaris, Wegener's granulomatosis, membranous nephropathy,
 amyotrophic lateral sclerosis, tabes dorsalis, giant cell
 arteritis/polymyalgia, pernicious anaemia, rapidly progressive
 glomerulonephritis, or fibrosing alveolitis. The infectious disease is
 chosen from bacterial, fungal, parasitic and viral lesion. The infectious
 disease is caused by a fungus chosen from Microsporium, Trichophyton,
 Epidermophyton, Sporothrix schenckii, Cryptococcus neoformans,
 Coccidioides immitis, Histoplasma capsulatum, Blastomycosis dermatitidis,
 and Candida albicans. The infectious disease is caused by a virus chosen
 from HIV, herpes virus, cytomegalovirus, rabies virus, influenza virus,
 hepatitis B virus, Sendai virus, feline leukemia virus, Reo virus, polio
 virus, human serum parvo-like virus, simian virus 40, respiratory
 syncytial virus, mouse mammary tumour virus, Varicella-Zoster virus,
 Dengue virus, rubella virus, measles virus, adenovirus, human T-cell
 leukemia viruses, Epstein-Barr virus, murine leukemia virus, mumps virus,
 vesicular stomatitis virus, Sindbis virus, lymphocytic choriomeningitis
 virus, wart virus and blue tongue virus. The infectious disease is caused
 by a bacterium chosen from Bacillus anthracis, Streptococcus agalactiae,
 Legionella pneumophila, Streptococcus pyogenes, Escherichia coli,
 Neisseria gonorrhoeae, Neisseria meningitidis, Pneumococcus Haemophilus
 influenzae B, Treponema pallidum, Lyme disease spirochetes, Pseudomonas
 aeruginosa, Mycobacterium leprae, Brucella abortus, Mycobacterium
 tuberculosis, and Tetanus toxin. The infectious disease is caused by a
 protozoa chosen from Plasmodium falciparum, Plasmodium vivax, Toxoplasma
 gondii, Trypanosoma rangeli, trypanosoma cruzi, Trypanosoma
 rhodesiense, Trypanosoma brucei, Schistosoma mansoni, Schistosoma
 japonicum, Babesia bovis, Elmeria tenella, Onchocerca volvulus,

CC Leishmania tropica, Trichinella spiralis, Onchocerca volvulus, Theileria
 CC parva, Taenia hydatigena, Taenia ovis, Taenia saginata, Echinococcus
 CC granulosus, and Mesocostoides corti. The infectious disease is caused by
 CC a mycoplasma chosen from Mycoplasma arthritidis, M. hyorhinis, M. orale,
 CC M. arginini, Acholeplasma laidlawii, M. salivarium and M. pneumoniae. The
 CC cancer is preferably chosen from carcinoembryonic antigen (CEA)-
 CC expressing tumour or a CD20-expressing malignancy. The present sequence
 CC represents a peptide used in the method of the invention.

XX Sequence 8 AA;

Query Match 80.6%; Score 43.5; DB 8; Length 8;
 Best Local Similarity 87.5%; Pred. No. 1.8e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GYWG-GYW 7
 |||||
 Db 1 GYWGKGYY 8

RESULT 9

AAAY76816
 ID AAY76816 standard; peptide; 9 AA.

XX AAY76816;

XX 28-APR-2000 (first entry)

XX Immunogenic peptide for bi-specific antibody recognition.

XX Immunogenic peptide; bi-specific antibody; diagnosis; immune response;
 XX diseased tissue identification; therapy.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "free amino acid group, protected amino acid
 FT group, chelating agent or a metal-chelate complex"

FT Misc-difference 3 /note= "D-form residue"

FT Misc-difference 4 /note= "D-form residue"

FT Misc-difference 6 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue; modified with free amino acid
 FT group, protected amino acid group, chelating agent or a
 FT metal-chelate complex"

FT Misc-difference 9 /note= "D-form residue"

FT Misc-difference 9 /note= "D-form residue"

XX WO966951-A2.

XX 29-DEC-1999.

XX 22-JUN-1999; 99WO-US013879.

XX 22-JUN-1998; 98US-0090142P.

XX 14-OCT-1998; 98US-0104156P.

XX (IMMU-) IMMUNOMEDICS INC.

XX Hansen HJ, Griffiths GL, Leung S, McBride WJ, Qu Z;

XX WPI; 2000-160561/14.

XX Bi-specific antibodies that bind specific target tissue and targeted
 XX conjugates.

XX Claim 21; Page 61; 76pp; English.

XX This sequence represents an immunogenic peptide for a bi-specific

CC antibody. The invention relates to a method of treating or identifying
 CC diseased tissues in a patient comprising administering a bi-specific
 CC antibody (or fragment) having at least 1 arm (A) that specifically
 CC binds a targeted tissue and at least 1 arm (B) that specifically binds a
 CC targetable conjugate. The methods and bi-specific antibodies and fusion
 CC proteins are useful for pre-targeting methods of diagnosis and therapy.
 CC It is advantageous to raise bi-specific antibodies against a targetable
 CC conjugate that is capable of carrying at least 1 diagnostic or
 CC therapeutic agent. The characteristics of the chelator, metal chelate
 CC complex, therapeutic agent or diagnostic agent can be varied to
 CC accommodate differing applications without raising new bi-specific
 CC antibodies for each new application. The targetable conjugate is selected
 CC to elicit sufficient immune responses and also for rapid in vivo
 CC clearance when used within the bi-specific antibody targeting method

XX Sequence 9 AA;

Query Match 80.6%; Score 43.5; DB 3; Length 9;

Best Local Similarity 87.5%; Pred. No. 1.8e+06;
 Matches 7; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GYWG-GYW 7
 |||||
 Db 2 GYWGKGYY 9

RESULT 10

AAO11208

ID AAO11208 standard; protein; 131 AA.

XX AAO11208;

XX 06-NOV-2001 (first entry)

XX Human polypeptide SEQ ID NO 25100.

XX Human; cytokine; cell proliferation; cell differentiation; gene therapy;
 XX vaccine; peptide therapy; stem cell growth factor; haematopoiesis;
 XX tissue growth factor; immunomodulatory; cancer; leukaemia;
 XX nervous system disorders; arthritis; inflammation.

XX Homo sapiens.

XX WO200164835-A2.

XX 07-SEP-2001.

XX 26-FEB-2001; 2001WO-US004927.

XX 28-FEB-2000; 2000US-00515126.

XX 18-MAY-2000; 2000US-00577409.

XX (HYSE-) HYSEQ INC.

XX Tang YT, Liu C, Drmanac RT;

XX WPI; 2001-514838/56.

XX N-PSDB; AAI91139.

XX Isolated nucleic acids and polypeptides, useful for preventing diagnosing
 XX and treating e.g. leukemia, inflammation and immune disorders.

XX Claim 20; SEQ ID NO 25100; 1399pp + Sequence Listing; English.

XX The invention relates to human polynucleotides (AAI79941-AAI93841) and
 CC the encoded proteins (AAO00010-AAO13910) that exhibit activity relating to
 CC cytokine, cell proliferation or cell differentiation or which may induce
 CC production of other cytokines in other cell populations. The
 CC polynucleotides and polypeptides are useful in gene therapy, vaccines or
 CC peptide therapy. The polypeptides have various cytokine-like activities,
 CC e.g. stem cell growth factor activity, haematopoiesis regulating
 CC activity, tissue growth factor activity, immunomodulatory activity and
 CC activin/inhibin activity and may be useful in the diagnosis and/or

CC treatment of cancer, leukaemia, nervous system disorders, arthritis and
CC inflammation. Note: The sequence data for this patent did not form part
CC of the printed specification, but was obtained in electronic format
CC directly from WIPO at ftp.wipo.int/pub/published_pct_sequences
XX
SQ Sequence 131 AA;

Query Match 79.6%; Score 43; DB 4; Length 131;
Best Local Similarity 83.3%; Pred. No. 83;
Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 2 YMGGYW 7
:|||||
Db 55 HWGGYW 60

RESULT 11

ADN49403
ID ADN49403 standard; protein; 118 AA.

XX AC ADN49403;

XX 01-JUL-2004 (first entry)

XX Human anti-IL-8 variable heavy chain (VH) antibody homologue protein.

XX Humanised antibody; therapy; diagnosis; autoimmune disease; cancer;
KW transplant rejection; infectious disease; inflammation;
KW variable heavy chain; VH; human.

XX Homo sapiens.

XX US2004067532-A1.

XX 08-APR-2004.

PF 11-JUN-2003; 2003US-00460595.

XX 12-AUG-2002; 2002US-0403296P.

XX (GENE-) GENETASTIX CORP.

XX Zhu L, Wei S, Hua SB;

XX WPI; 2004-355295/33.

XX Screening library of humanized antibody useful in therapeutics, by
PT expressing target protein in yeast expressing humanized antibodies, and
PT selecting yeast in which reporter gene is expressed on activation by
PT binding of antibody to protein.

XX Disclosure; SEQ ID NO 8; 38pp; English.

XX The present invention provides compositions, methods and kits for
CC efficiently generating and screening humanised antibody with high
CC affinity against a specific antigen. The invention is useful in
CC therapeutics and the diagnosis of diseases such as autoimmune diseases,
CC cancer, transplant rejection, infectious diseases and inflammation. The
CC present sequence is human anti-IL-8 variable heavy chain (VH) antibody
CC homologue protein.

XX SQ Sequence 118 AA;

Query Match 77.8%; Score 42; DB 8; Length 118;
Best Local Similarity 83.3%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YMGGYW 7

Db 101 YWGNW 106
|||||

RESULT 12

ABU39959
ID ABU39959 standard; protein; 211 AA.

XX AC ABU39959;

XX 19-JUN-2003 (first entry)

XX Protein encoded by Prokaryotic essential gene #25486.

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX Pseudomonas putida.

XX WO200277183-A2.

XX 03-OCT-2002.

XX 21-MAR-2002; 2002WO-US009107.

XX 21-MAR-2001; 2001US-00815242.

XX 06-SEP-2001; 2001US-00948993.

XX 25-OCT-2001; 2001US-0342923P.

XX 08-FEB-2002; 2002US-00072851.

XX 06-MAR-2002; 2002US-0362699P.

XX (ELIT-) ELITRA PHARM INC.

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;

XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

XX WPI; 2003-029926/02.

XX N-PSDB; ACA43829.

XX New antisense nucleic acids, useful for identifying proteins or screening
PT for homologous nucleic acids required for cellular proliferation to
PT isolate candidate molecules for rational drug discovery programs.

XX Claim 25; SEQ ID NO 67883; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX Sequence 211 AA;

Query Match 77.8%; Score 42; DB 6; Length 211;
 Best Local Similarity 71.4%; Pred. No. 1.8e+02;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 1 GYWGYYW 7
 36 GFWRGYW 42

Db

RESULT 13
 AAW09436
 ID AAW09436 standard; protein; 243 AA.
 XX
 AC AAW09436;
 XX
 XX
 DT 23-SEP-1997 (first entry)
 XX
 DE Anti-CD19 antibody BLV3 single chain variable region.
 XX
 KW CD19; antibody; variable region; B43; SJ25C1; BLV3; targeting;
 KW cytotoxin conjugate; cancer treatment; leukaemia; B cell lymphoma;
 KW variable; heavy; light chain.
 XX
 OS Synthetic.

XX Key Location/Qualifiers
 FH Domain 1..114
 FT /label= heavy_chain
 FT 1..26
 FT /label= FR1
 FT 27..31
 FT /label= CDR1
 FT 32..45
 FT /label= FR2
 FT 46..62
 FT /label= CDR2
 FT 63..94
 FT /label= FR3
 FT 95..114
 FT /label= CDR3
 FT 115..129
 FT /label= linker
 FT 130..243
 FT /label= light_chain
 FT 130..152
 FT /label= FR1
 FT 153..167
 FT /label= CDR1
 FT 168..182
 FT /label= FR2
 FT 183..189
 FT /label= CDR2
 FT 190..221
 FT /label= FR3
 FT 222..243
 FT /label= CDR3

XX WO9636360-A1.
 XX
 XX 21-NOV-1996.
 XX
 XX 15-MAY-1996; 96WO-US006941.
 XX
 XX 17-MAY-1995; 95US-00443408.
 XX
 PA (MINU) UNIV MINNESOTA.
 XX
 XX Kersey JH, Bejcek BE, Wang D, Uckun FM;
 PI 1997-042642/04.
 XX N-PSDB; AAT47734.
 XX
 XX DNA encoding a single chain variable region polypeptide which binds CD19

PT - used in the treatment of leukaemia and B-cell lymphoma.
 XX
 XX Claim 2; Page 105-106; 143pp; English.
 XX
 CC AAW09436 represents a single chain variable region polypeptide (heavy
 CC chain-linker-light chain) of anti-CD19 antibody BLV3. Single chain
 CC polypeptides were conjugated to at least one cytotoxic agent and used to
 CC target the agent(s) to cancerous cells expressing CD19. The
 CC immunconjugates can be used treatment of cancer, especially leukaemia
 CC and B-cell lymphoma
 XX
 SQ Sequence 243 AA;
 XX
 Query Match 77.8%; Score 42; DB 2; Length 243;
 Best Local Similarity 83.3%; Pred. No. 2e+02;
 Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YWGYYW 7
 97 YWGYW 102

Db

RESULT 14
 ABM67619
 ID ABM67619 standard; protein; 447 AA.
 XX
 AC ABM67619;
 XX
 DT 20-NOV-2003 (first entry)
 XX
 DE Photorhabdus luminescens protein sequence #716.
 XX
 KW Antibacterial; fungicide; insecticide; polymorphism; genetic analysis;
 KW detection; food; gene expression; plant; animal; microorganism; toxin;
 KW antibiotic; biopesticide; virulence factor; disease model; plague;
 KW whooping cough.
 XX
 OS Photorhabdus luminescens.
 XX
 FN WO200294867-A2.
 XX
 PD 28-NOV-2002.
 XX
 PF 07-FEB-2002; 2002WO-IB003040.
 XX
 PR 07-FEB-2001; 2001FR-00001659.
 XX
 PA (INSP) INST PASTEUR.
 XX (CNRS) CNRS CENT NAT RECH SCI.
 XX
 PI Duchaud E, Taourit S, Glaser P, Frangeul L, Kunst F, Danchin A;
 PI Buchrieser C;
 XX
 DR WPI; 2003-148459/14.
 XX
 XX Genomic sequence of Photorhabdus luminescens and encoded polypeptides,
 PT useful e.g. as therapeutic antimicrobials and agricultural pesticides.
 XX
 XX Claim 2; SEQ ID NO 716; 1205pp; French.
 XX
 CC The invention relates to the isolation of genes and their encoded
 CC proteins from Photorhabdus luminescens. The isolated sequences are
 CC sources of probes and primers for detecting the genome of P. luminescens
 CC and related species; to study polymorphisms; for gene analysis and for
 CC detection/amplification of the genes. Antibodies (Ab) raised against the
 CC polypeptides encoded by the genes are used for detection/identification
 CC of P. luminescens, e.g. in foods. The genes, proteins, Ab and cells that
 CC carry a gene-containing vector are used to select compounds that
 CC modulate, regulate, induce or inhibit expression of the genes in plants,
 CC animals or microorganisms other than P. luminescens and are able to alter
 CC response or sensitivity to toxins and antibiotics produced by P.
 CC luminescens. Cells transformed to express the genes are useful for
 CC recombinant production of the proteins, particularly toxins and

CC antibacterials useful as insecticides, bactericides and fungicides. The
CC genes, proteins, vectors containing the genes and Ab are also useful
CC therapeutically (to treat microbial infection by bacteria or fungi that
CC are sensitive to P. luminescens-encoded toxins or antibiotics) and as
CC biopesticides. Other uses of the genes and the proteins are as virulence
CC factors and for identifying targets of human diseases for which P.
CC luminescens is a model (particularly plague and whooping cough). This
CC sequence represents one of the isolated P. luminescens proteins
XX
SQ Sequence 447 AA;
Query Match 77.8%; Score 42; DB 6; Length 447;
Best Local Similarity 71.4%; Pred. NO. 3.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;
QY 1 GYWGYYW 7
|:|:|:|
Db 147 GFWWGYW 153
RESULT 15
AAU90607
ID AAU90607 standard; peptide; 20 AA.
XX
AC AAU90607;
XX
DT 18-JUN-2002 (first entry)
XX
DE Insulin/insulin-like growth factor receptor-binding peptide #2563.
XX
KW Cytostatic; antidiabetic; neuroprotective; cerebroprotective;
KW ophthalmological; insulin; receptor; gene therapy; diabetes;
KW insulin-like growth factor-1; IGF-1; tumour; prostate; breast;
KW diabetic retinopathy; neurological diseases; stroke; diabetic neuropathy.
XX
OS Synthetic.
XX
PN WO200172771-A2.
XX
PD 04-OCT-2001.
XX
PF 29-MAR-2000; 2000WO-US008528.
XX
PR 29-MAR-2000; 2000WO-US008528.
XX
PA (DGIB-) DGI BIOTECHNOLOGIES LLC.
PA (NOVO) NOVO NORDISK AS.
XX
PI Beasley J, Blume AJ, Schaeffer L, Pillutla R, Brandt J;
PI Brissette R, Spetzler J, Cheng W, Ostergaard S, Manddecki WS;
PI Hansen PH, Ravera M, Hsiao K;
XX
DR WPI; 2002-025774/03.
XX
PT Modulating insulin activity in mammalian cells, for treating e.g.
PT diabetes and tumors, comprises using peptides that bind to insulin or
PT insulin-like growth factor receptors.
XX
PS Example 9; Fig 10A-2; 390pp; English.
XX
CC The invention relates to a method of modulating insulin activity in
CC mammalian cells by administering a peptide that binds the insulin
CC receptor (IR). A composition containing a peptide, optionally expressed
CC from gene therapy vectors, that binds to Site 1 of IR and an insulin
CC agonist are useful for treating diabetes. Also, peptides that are
CC antagonists of the insulin-like growth factor-1 (IGF-1) receptor are
CC useful for treating insulin-like growth factor (IGF)-sensitive tumours
CC (e.g. of prostate and breast) and diabetic retinopathy, while IGF-1
CC receptor agonists are useful for treating neurological diseases,
CC including stroke and diabetic neuropathy. The peptides are also useful in
CC screening for compounds that bind to IR or IGF-1 receptor, potential
CC therapeutics and research reagents. AAU8034-AAU90957 represent IR and/or
CC IGF-1 receptor-binding peptides and related amino acid sequences of the

CC invention
XX Sequence 20 AA;
SQ
Query Match 75.9%; Score 41; DB 5; Length 20;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 WGGYW 7
|:|:|:|
Db 13 WGGYW 17
Search completed: April 18, 2005, 19:50:23
Job time : 73.7907 secs

THIS PAGE BLANK (USPTO)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 19:52:14 ; Search time 51.2791 Seconds
(without alignments)
45.371 Million cell updates/sec

Title: SEQ1

Perfect score: 54

Sequence: 1 gywgyw 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications_AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09D_PUBCOMB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match %	Length	ID	Description
1	54	100.0	440	15	US-10-282-122A-48322 Sequence 48322, A
2	48	88.9	296	14	US-10-156-761-9632 Sequence 9632, Ap
3	46	85.2	247	15	US-10-424-599-184877 Sequence 184877, A
4	45	83.3	141	16	US-10-437-963-143884 Sequence 143884, A
5	42	77.8	118	15	US-10-460-595-8 Sequence 8, Appli
6	42	77.8	211	15	US-10-282-122A-67883 Sequence 67883, A
7	41	75.9	20	10	US-09-962-756-1075 Sequence 1075, Ap
8	41	75.9	20	15	US-10-253-471-1075 Sequence 1075, Ap
9	41	75.9	20	15	US-10-253-493-1075 Sequence 1075, Ap
10	41	75.9	26	10	US-09-962-756-1516 Sequence 1516, Ap
11	41	75.9	26	15	US-10-253-471-1516 Sequence 1516, Ap
12	41	75.9	26	15	US-10-253-493-1516 Sequence 1516, Ap
13	41	75.9	40	10	US-09-962-756-1113 Sequence 1113, Ap

14	41	75.9	40	15	US-10-253-471-1113 Sequence 1113, Ap
15	41	75.9	40	15	US-10-253-493-1113 Sequence 1113, Ap
16	41	75.9	54	15	US-10-424-599-246894 Sequence 246894, A
17	41	75.9	105	10	US-09-933-767-745 Sequence 745, App
18	41	75.9	105	14	US-10-004-860-745 Sequence 745, App
19	41	75.9	105	14	US-10-023-282-745 Sequence 745, App
20	41	75.9	396	15	US-10-282-122A-47856 Sequence 47856, A
21	41	75.9	396	15	US-10-282-122A-50261 Sequence 50261, A
22	41	75.9	400	15	US-10-369-493-1129 Sequence 1129, Ap
23	41	75.9	403	15	US-10-282-122A-49552 Sequence 49552, A
24	40	74.1	141	15	US-10-424-599-190607 Sequence 190607, A
25	40	74.1	347	14	US-10-156-761-7823 Sequence 7823, Ap
26	40	74.1	399	15	US-10-424-599-151557 Sequence 151557, A
27	40	74.1	543	15	US-10-345-820-99 Sequence 99, Appl
28	40	74.1	551	16	US-10-437-963-181262 Sequence 181262, A
29	40	74.1	552	15	US-10-345-820-95 Sequence 95, Appl
30	40	74.1	552	15	US-10-345-820-102 Sequence 102, App
31	40	74.1	552	15	US-10-345-820-105 Sequence 105, App
32	40	74.1	552	15	US-10-345-820-108 Sequence 108, App
33	40	74.1	552	15	US-10-345-820-111 Sequence 111, App
34	40	74.1	552	15	US-10-345-820-115 Sequence 115, App
35	40	74.1	552	15	US-10-345-820-118 Sequence 118, App
36	40	74.1	564	16	US-10-437-963-180291 Sequence 180291, A
37	40	74.1	576	16	US-10-437-963-122848 Sequence 122848, A
38	40	74.1	585	14	US-10-342-224-32 Sequence 32, Appl
39	40	74.1	587	16	US-10-767-701-45092 Sequence 45092, A
40	40	74.1	591	16	US-10-437-963-181854 Sequence 181854, A
41	40	74.1	602	15	US-10-369-493-22275 Sequence 22275, A
42	40	74.1	1039	15	US-10-369-493-6849 Sequence 6849, Ap
43	40	74.1	1827	9	US-09-819-247-2 Sequence 2, Appli
44	40	74.1	1827	15	US-10-428-225-2 Sequence 2, Appli
45	40	74.1	1827	16	US-10-751-699-14 Sequence 14, Appl

ALIGNMENTS

RESULT 1

US-10-282-122A-48322
; Sequence 48322, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931

; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48322
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Bacteroides fragilis
US-10-282-122A-48322

Query Match 100.0%; Score 54; DB 15; Length 440;
Best Local Similarity 100.0%; Pred. No. 7.4; Indels 0; Gaps 0;
Matches 7; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GYWGYYW 7
|||||
Db 212 GYWGYYW 218

RESULT 2
US-10-156-761-9632
; Sequence 9632, Application US/10156761
; Publication No. US20030119018A1
; GENERAL INFORMATION:
; APPLICANT: OMURA, SATOSHI
; APPLICANT: IKEDA, HARUO
; APPLICANT: ISHIKAWA, HIROSHI
; APPLICANT: HORIKAWA, TADAYOSHI
; APPLICANT: SHIBA, TADAYOSHI
; APPLICANT: SAKAKI, YOSHIYUKI
; APPLICANT: HATTORI, MASAHIRA
; TITLE OF INVENTION: NOVEL POLYNUCLEOTIDES
; FILE REFERENCE: 249-262
; CURRENT APPLICATION NUMBER: US/10/156,761
; CURRENT FILING DATE: 2002-05-29
; PRIOR APPLICATION NUMBER: JP 2001-204089
; PRIOR FILING DATE: 2001-05-30
; PRIOR APPLICATION NUMBER: JP 2001-272697
; PRIOR FILING DATE: 2001-08-02
; NUMBER OF SEQ ID NOS: 15109
; SEQ ID NO 9632
; LENGTH: 296
; TYPE: PRT
; ORGANISM: Streptomyces avermitilis
US-10-156-761-9632

Query Match 88.9%; Score 48; DB 14; Length 296;
Best Local Similarity 85.7%; Pred. No. 36;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWGYYW 7
|||||
Db 198 GYWGYYW 204

RESULT 3
US-10-424-599-184877
; Sequence 184877, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 184877
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(247)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_13795C.1.pap
US-10-424-599-184877

Query Match 85.2%; Score 46; DB 15; Length 247;
Best Local Similarity 85.7%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWGYYW 7
|||||
Db 109 GYWGYYW 115

RESULT 4
US-10-437-963-143884
; Sequence 143884, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated with
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 143884
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_4474C.1.pap
US-10-437-963-143884

Query Match 83.3%; Score 45; DB 16; Length 141;
Best Local Similarity 85.7%; Pred. No. 52;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWGYYW 7
|||||
Db 4 GYWGYYW 10

RESULT 5
US-10-460-595-8
; Sequence 8, Application US/10460595
; Publication No. US20040067532A1
; GENERAL INFORMATION:
; APPLICANT: Zhu, Li
; APPLICANT: Wei, Shuanghong
; APPLICANT: Hua, Shaobing B
; TITLE OF INVENTION: HIGH THROUGHPUT GENERATION AND AFFINITY MATURATION OF HUMANIZED
; FILE REFERENCE: 25636-730
; CURRENT APPLICATION NUMBER: US/10/460,595
; CURRENT FILING DATE: 2003-06-11
; PRIOR APPLICATION NUMBER: US 60/403,296
; PRIOR FILING DATE: 2002-08-12
; NUMBER OF SEQ ID NOS: 9

```
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 118
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-460-595-8

Query Match      77.8%; Score 42; DB 15; Length 118;
Best Local Similarity 83.3%; Pred. No. 1.2e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 YWGGYW 7
DB      101 YWGNYW 106

RESULT 6
US-10-282-122A-67883
; Sequence 67883, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 67883
; LENGTH: 211
; TYPE: PRT
; ORGANISM: Pseudomonas putida
US-10-282-122A-67883

Query Match      77.8%; Score 42; DB 15; Length 211;
Best Local Similarity 71.4%; Pred. No. 1.8e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYWGYYW 7
DB      36 GFWRGYW 42

; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-962-756-1075
; Sequence 1075, Application US/09962756
; Publication No. US20030195147A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA
; APPLICANT: BRISSETTE, RENEE
; APPLICANT: BLUME, ARTHUR J.
; APPLICANT: SCHAEFER, LAUGE
; APPLICANT: BRANDT, JAKOB
; APPLICANT: GOLDSTEIN, NEIL I.
; APPLICANT: SPETZLER, JANE
; APPLICANT: OSTERGAARD, SOREN
; APPLICANT: HANSEN, PER HERTZ
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4051US1
; CURRENT APPLICATION NUMBER: US/09/962,756
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1075
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-09-962-756-1075

Query Match      75.9%; Score 41; DB 10; Length 20;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 WGGYW 7
DB      13 WGGYW 17

RESULT 8
US-10-253-471-1075
; Sequence 1075, Application US/10253471
; Publication No. US20030236190A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4057
; CURRENT APPLICATION NUMBER: US/10/253,471
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1075
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
US-10-253-471-1075

Query Match      75.9%; Score 41; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 40;
```

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 3 WGGYW 7
| | | | |
Db 13 WGGYW 17

RESULT 9
US-10-253-493-1075
; Sequence 1075, Application US/10253493
; Publication No. US20040023887A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4056
; CURRENT APPLICATION NUMBER: US/10/253,493
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1075
; LENGTH: 20
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-253-493-1075

Query Match 75.9%; Score 41; DB 15; Length 20;
Best Local Similarity 100.0%; Pred. No. 40;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WGGYW 7
| | | | |
Db 13 WGGYW 17

RESULT 10
US-09-962-756-1516
; Sequence 1516, Application US/09962756
; Publication No. US20030195147A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA
; APPLICANT: BRISSETTE, RENEE
; APPLICANT: BLUME, ARTHUR J.
; APPLICANT: SCHAFER, LAUGE
; APPLICANT: BRANDT, JAKOB
; APPLICANT: GOLDSTEIN, NEIL I.
; APPLICANT: SPETZLER, JANE
; APPLICANT: OSTERGAARD, SOREN
; APPLICANT: HANSEN, PER HERTZ
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4051US1
; CURRENT APPLICATION NUMBER: US/09/962,756
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1516
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic

; OTHER INFORMATION: peptide
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (19)
; OTHER INFORMATION: Unknown amino acid
US-09-962-756-1516

Query Match 75.9%; Score 41; DB 10; Length 26;
Best Local Similarity 71.4%; Pred. No. 49;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYGGYW 7
| : | | | |
Db 20 GFWGGSW 26

RESULT 11
US-10-253-471-1516
; Sequence 1516, Application US/10253471
; Publication No. US20030236190A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4057
; CURRENT APPLICATION NUMBER: US/10/253,471
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1516
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
US-10-253-471-1516

; OTHER INFORMATION: Unknown amino acid
; LOCATION: (19)
; OTHER INFORMATION: peptide
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (19)
; OTHER INFORMATION: Unknown amino acid
US-10-253-471-1516

Query Match 75.9%; Score 41; DB 15; Length 26;
Best Local Similarity 71.4%; Pred. No. 49;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYGGYW 7
| : | | | |
Db 20 GFWGGSW 26

RESULT 12
US-10-253-493-1516
; Sequence 1516, Application US/10253493
; Publication No. US20040023887A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4056
; CURRENT APPLICATION NUMBER: US/10/253,493
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227

```

; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1516
; LENGTH: 26
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (19)
; OTHER INFORMATION: Unknown amino acid
US-10-253-493-1516

Query Match 75.9%; Score 41; DB 15; Length 26;
Best Local Similarity 71.4%; Pred. No. 49;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYGGYW 7
|:|||||
Db 20 GFWGSW 26

RESULT 13
US-09-962-756-1113
; Sequence 1113, Application US/09962756
; Publication No. US20030195147A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA
; APPLICANT: BRISSETTE, RENEE
; APPLICANT: BLUME, ARTHUR J.
; APPLICANT: SCHAFER, LAUGE
; APPLICANT: BRANDT, JAKOB
; APPLICANT: GOLDSTEIN, NEIL I.
; APPLICANT: SPETZLER, JANE
; APPLICANT: OSTERGAARD, SOREN
; APPLICANT: HANSEN, PER HERTZ
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4051US1
; CURRENT APPLICATION NUMBER: US/09/962,756
; CURRENT FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1113
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (31)
; OTHER INFORMATION: Unknown amino acid
US-09-962-756-1113

Query Match 75.9%; Score 41; DB 10; Length 40;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WGGYW 7
|:|||||
Db 33 WGGYW 37

RESULT 14
US-10-253-471-1113
; Sequence 1113, Application US/10253471
; Publication No. US20030236190A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4056
; CURRENT APPLICATION NUMBER: US/10/253,493
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1113
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (31)
; OTHER INFORMATION: Unknown amino acid
US-10-253-493-1113

Query Match 75.9%; Score 41; DB 15; Length 40;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WGGYW 7
|:|||||
Db 33 WGGYW 37

RESULT 15
US-10-253-493-1113
; Sequence 1113, Application US/10253493
; Publication No. US20040023887A1
; GENERAL INFORMATION:
; APPLICANT: PILLUTLA, RENUKA et al.
; TITLE OF INVENTION: INSULIN AND IGF-1 RECEPTOR AGONISTS AND ANTAGONISTS
; FILE REFERENCE: 1878-4056
; CURRENT APPLICATION NUMBER: US/10/253,493
; CURRENT FILING DATE: 2002-09-24
; PRIOR APPLICATION NUMBER: 09/962,756
; PRIOR FILING DATE: 2001-09-24
; PRIOR APPLICATION NUMBER: 09/538,038
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: 09/146,127
; PRIOR FILING DATE: 1998-09-02
; NUMBER OF SEQ ID NOS: 2227
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1113
; LENGTH: 40
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence: Synthetic
; OTHER INFORMATION: peptide
; FEATURE:
; NAME/KEY: MOD RES
; LOCATION: (31)
; OTHER INFORMATION: Unknown amino acid
US-10-253-493-1113

Query Match 75.9%; Score 41; DB 15; Length 40;
Best Local Similarity 100.0%; Pred. No. 68;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 WGGYW 7
|:|||||
Db 33 WGGYW 37

```

Tue Apr 19 06:58:37 2005

seq1.rapb

Page 6

Db 33 WGGYW 37

Search completed: April 18, 2005, 20:29:25
Job time : 52.2791 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 19:38:48 ; Search time 19.5349 Seconds
(without alignments)
26.749 Million cell updates/sec

Title: SEQ1
Perfect score: 54
Sequence: 1 GYWGYYW 7

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA: +
1: /cgn2_6/protdata/1/iaa/5A COMB.pep.*
2: /cgn2_6/protdata/1/iaa/5B COMB.pep.*
3: /cgn2_6/protdata/1/iaa/6A COMB.pep.*
4: /cgn2_6/protdata/1/iaa/6B COMB.pep.*
5: /cgn2_6/protdata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/protdata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	54	100.0	339	4	US-09-252-991A-26841
2	45	83.3	500	4	US-09-252-991A-21214
3	41	75.9	105	4	US-09-205-258-745
4	41	75.9	571	4	US-09-302-540-14983
5	40	74.1	585	1	US-08-212-188-4
6	40	74.1	585	3	US-08-970-725-4
7	40	74.1	585	5	PCT-US95-02708-4
8	40	74.1	1827	4	US-09-443-780C-14
9	40	74.1	1827	4	US-09-079-723-179
10	39	72.2	1621	3	US-08-972-927-3
11	39	72.2	1622	3	US-08-972-927-6
12	38	70.4	12	1	US-08-241-054-99
13	38	70.4	12	1	US-08-439-817-79
14	38	70.4	12	1	US-08-485-508-99
15	38	70.4	86	4	US-09-248-796A-22524
16	38	70.4	212	4	US-09-543-681A-5065
17	38	70.4	470	4	US-09-252-991A-23310
18	38	70.4	638	2	US-08-846-762-95
19	37	68.5	80	4	US-09-252-991A-22806
20	37	68.5	150	4	US-09-371-056-16
21	37	68.5	154	4	US-09-502-540-16787
22	37	68.5	231	3	US-08-448-489-19
23	37	68.5	231	4	US-09-689-730-19
24	37	68.5	275	4	US-08-371-056-10
25	37	68.5	335	1	US-08-118-270-11
26	37	68.5	355	5	PCT-US93-08528-11
27	37	68.5	489	1	US-08-489-733-6

28	37	68.5	489	2	US-08-993-581B-6	Sequence 6, Appli
29	37	68.5	491	1	US-08-489-733-5	Sequence 5, Appli
30	37	68.5	491	2	US-08-993-581B-5	Sequence 5, Appli
31	37	68.5	515	4	US-09-248-796A-19787	Sequence 19787, A
32	37	68.5	822	4	US-09-248-796A-14728	Sequence 14728, A
33	36.5	67.6	359	4	US-09-543-681A-5995	Sequence 5995, Ap
34	36.5	67.6	359	4	US-09-489-039A-13566	Sequence 13566, A
35	36	66.7	84	4	US-09-513-999C-5403	Sequence 5403, Ap
36	36	66.7	136	4	US-09-270-767-33627	Sequence 33627, A
37	36	66.7	136	4	US-09-270-767-48844	Sequence 48844, A
38	36	66.7	160	2	US-08-162-402B-14	Sequence 14, Appli
39	36	66.7	218	1	US-07-607-538C-4	Sequence 4, Appli
40	36	66.7	218	2	US-08-162-402B-4	Sequence 4, Appli
41	36	66.7	218	4	US-09-364-185-4	Sequence 4, Appli
42	36	66.7	246	1	US-08-271-354-9	Sequence 9, Appli
43	36	66.7	246	2	US-08-565-861-9	Sequence 9, Appli
44	36	66.7	246	5	PCT-US94-07658-9	Sequence 9, Appli
45	36	66.7	250	4	US-09-702-705-1677	Sequence 1677, Ap

ALIGNMENTS

RESULT 1

US-09-252-991A-26841

; Sequence 26841, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 26841

; LENGTH: 339

; TYPE: PRT

; ORGANISM: Pseudomonas aeruginosa

US-09-252-991A-26841

Query Match 100.0%; Score 54; DB 4; Length 339;

Best Local Similarity 100.0%; Pred. No. 1.4; Indels 0; Gaps 0;

Matches 7; Conservative 0; Mismatches 0;

QY 1 GYWGYYW 7

DB 261 GYWGYYW 267

RESULT 2

US-09-252-991A-21214

; Sequence 21214, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

; SEQ ID NO 21214

; LENGTH: 500

; TYPE: PRT

```
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-21214

Query Match      83.3%; Score 45; DB 4; Length 500;
Best Local Similarity 85.7%; Pred. No. 38;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYGGYW 7
Db      331 GYGGRW 337

RESULT 3
US-09-205-258-745
; Sequence 745, Application US/09205258
; Patent No. 6525174
; GENERAL INFORMATION:
; APPLICANT: Young et al.
; TITLE OF INVENTION: 207 Human Secreted Proteins
; CURRENT APPLICATION NUMBER: US/09/205,258
; CURRENT FILING DATE: 1998-12-04
; EARLIER APPLICATION NUMBER: PCT/US98/11422
; EARLIER FILING DATE: 1998-06-04
; EARLIER APPLICATION NUMBER: 60/048,885
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,375
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,881
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,880
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,896
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,020
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,876
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,895
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,884
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,894
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,971
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,964
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,882
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,899
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,893
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,900
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,901
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,892
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,915
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,019
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,970
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,972
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,916
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,373
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,875

; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/049,374
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,917
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,949
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,974
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,883
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,897
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,898
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,962
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,963
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,877
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/048,878
; EARLIER FILING DATE: 1997-06-06
; EARLIER APPLICATION NUMBER: 60/070,923
; EARLIER FILING DATE: 1997-12-18
; EARLIER APPLICATION NUMBER: 60/092,921
; EARLIER FILING DATE: 1998-07-15
; EARLIER APPLICATION NUMBER: 60/094,657
; EARLIER FILING DATE: 1998-07-30
; NUMBER OF SEQ ID NOS: 1227
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 745
; LENGTH: 105
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-205-258-745

Query Match      75.9%; Score 41; DB 4; Length 105;
Best Local Similarity 71.4%; Pred. No. 33;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY      1 GYGGYW 7
Db      11 GYGGRW 17

RESULT 4
US-09-902-540-14983
; Sequence 14983, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 14983
; LENGTH: 571
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
US-09-902-540-14983

Query Match      75.9%; Score 41; DB 4; Length 571;
Best Local Similarity 100.0%; Pred. No. 1.6e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      3 WGGYW 7
```

```
Db      134 WGGYW 138
|||||
RESULT 5
US-08-212-188-4
; Sequence 4, Application US/08212188
; Patent No. 5689039
; GENERAL INFORMATION:
; APPLICANT: BECKER, JEFFREY M.
; APPLICANT: STACEY, GARY
; TITLE OF INVENTION: PLANT PEPTIDE TRANSPORT GENE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWREY & SIMON
; STREET: 1299 PENNSYLVANIA AVE., N.W.
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/212,188
; FILING DATE:
; CLASSIFICATION: 435
; ATTORNEY/AGENT INFORMATION:
; NAME: AUERBACH, JEFFREY I
; REGISTRATION NUMBER: 32,680
; REFERENCE/DOCKET NUMBER: 7493-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 383-7451
; TELEFAX: (202) 383-6610
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; CLONE: ATPR2Bp
; US-08-212-188-4

Query Match      74.1%; Score 40; DB 1; Length 585;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 YWGGYW 7
Db      109 YWGRYW 114
|||||
RESULT 6
US-08-970-725-4
; Sequence 4, Application US/08970725
; Patent No. 6080542
; GENERAL INFORMATION:
; APPLICANT: Becker, Jeffrey M.
; APPLICANT: Stacey, Gary
; TITLE OF INVENTION: PLANT PEPTIDE TRANSPORT GENE
; NUMBER OF SEQUENCES: 14
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: WEISER & ASSOCIATES
; STREET: 230 South Fifteenth Street, Suite 500
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA

Query Match      74.1%; Score 40; DB 1; Length 585;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 YWGGYW 7
Db      109 YWGRYW 114
|||||
RESULT 7
PCT-US95-02708-4
; Sequence 4, Application PC/TUS9502708
; GENERAL INFORMATION:
; APPLICANT: BECKER, JEFFREY M.
; APPLICANT: STACEY, GARY
; TITLE OF INVENTION: PLANT PEPTIDE TRANSPORT GENE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWREY & SIMON
; STREET: 1299 PENNSYLVANIA AVE., N.W.
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02708
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: AUERBACH, JEFFREY I
; REGISTRATION NUMBER: 32,680
; REFERENCE/DOCKET NUMBER: 7493-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 383-7451
; TELEFAX: (202) 383-6610
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: NO
; ORIGINAL SOURCE:
; ORGANISM: Arabidopsis thaliana
; IMMEDIATE SOURCE:
; CLONE: ATPR2Bp
; US-08-212-188-4

Query Match      74.1%; Score 40; DB 3; Length 585;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 YWGGYW 7
Db      109 YWGRYW 114
|||||
PCT-US95-02708-4
; Sequence 4, Application PC/TUS9502708
; GENERAL INFORMATION:
; APPLICANT: BECKER, JEFFREY M.
; APPLICANT: STACEY, GARY
; TITLE OF INVENTION: PLANT PEPTIDE TRANSPORT GENE
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HOWREY & SIMON
; STREET: 1299 PENNSYLVANIA AVE., N.W.
; CITY: WASHINGTON
; STATE: DC
; COUNTRY: US
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/02708
; FILING DATE:
; CLASSIFICATION:
; ATTORNEY/AGENT INFORMATION:
; NAME: AUERBACH, JEFFREY I
; REGISTRATION NUMBER: 32,680
; REFERENCE/DOCKET NUMBER: 7493-006
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (202) 383-7451
; TELEFAX: (202) 383-6610
; INFORMATION FOR SEQ ID NO: 4:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 585 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
```

MOLECULE TYPE: protein
HYPOTHETICAL: NO
ORIGINAL SOURCE:
ORGANISM: Arabidopsis thaliana
IMMEDIATE SOURCE:
CLONE: ATPT2BP
PCT-US95-02708-4

Query Match 74.1%; Score 40; DB 5; Length 585;
Best Local Similarity 83.3%; Pred. No. 2.3e+02;
Matches 5; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YWGGYW 7
| | | | |
DB 109 YWGRYW 114

RESULT 8
US-09-443-780C-14
Sequence 14, Application US/09443780C
Patent No. 6699973
GENERAL INFORMATION:
APPLICANT: O'Mahony, Daniel J
APPLICANT: Seveso, Michela
TITLE OF INVENTION: Antibodies To Peptides That Target GIT Receptors And Related Meth
FILE REFERENCE: E1067/20037
CURRENT APPLICATION NUMBER: US/09/443,780C
CURRENT FILING DATE: 1999-11-19
PRIOR APPLICATION NUMBER: US 60/109,036
PRIOR FILING DATE: 1998-11-19
NUMBER OF SEQ ID NOS: 84
SOFTWARE: PatentIn version 3.1
SEQ ID NO 14
LENGTH: 1827
TYPE: PRT
ORGANISM: Artificial
FEATURE:
OTHER INFORMATION: hSI receptor
US-09-443-780C-14

Query Match 74.1%; Score 40; DB 4; Length 1827;
Best Local Similarity 71.4%; Pred. No. 6.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWGYYW 7
| | | | |
DB 1491 GRWGHW 1497

RESULT 9
US-09-079-723-179
Sequence 179, Application US/09079723
Patent No. 6703362
GENERAL INFORMATION:
APPLICANT: Alvarez, Vernon L.
APPLICANT: O'Mahony, Daniel J.
APPLICANT: Lambkin, Imelda J.
APPLICANT: Singleton, Judith
APPLICANT: Patterson, Catherine A.
APPLICANT: Cagney, Gerard W.
APPLICANT: Belinka, Benjamin A.
APPLICANT: Carter, John M.
TITLE OF INVENTION: RANDOM PEPTIDES THAT BIND TO GASTRO-
NUMBER OF INVENTION: 265
NUMBER OF SEQUENCES: 265
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds LLP
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036
COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/079,723
FILING DATE:
CLASSIFICATION:
ATTORNEY/AGENT INFORMATION:
NAME: Misrock, S. Lealie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 1101-219
TELECOMMUNICATION INFORMATION:
TELEPHONE: 212-790-9090
TELEFAX: 212-869-9741
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 179:
SEQUENCE CHARACTERISTICS:
LENGTH: 1827 amino acids
TYPE: amino acid
STRANDEDNESS:
TOPOLOGY: unknown
MOLECULE TYPE: peptide
US-09-079-723-179

Query Match 74.1%; Score 40; DB 4; Length 1827;
Best Local Similarity 71.4%; Pred. No. 6.5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 GYWGYYW 7
| | | | |
DB 1491 GRWGHW 1497

RESULT 10
US-08-972-927-3
Sequence 3, Application US/08972927
Patent No. 6166290
GENERAL INFORMATION:
APPLICANT: Rea, Philip A
APPLICANT: Lu, Yu-Ping
APPLICANT: Li, Ze-Sheng
TITLE OF INVENTION: GLUTATHIONE-S-CONJUGATE TRANSPORT IN
NUMBER OF SEQUENCES: 20
CORRESPONDENCE ADDRESS:
ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
STREET: One Commerce Square, 2005 Market Street, 22nd
CITY: Philadelphia
STATE: Pennsylvania
COUNTRY: US
ZIP: 19103-7086
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/972,927
FILING DATE: 18-NOV-1997
CLASSIFICATION: 800
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/031,040
FILING DATE: 18-NOV-1996
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 60/061,328
FILING DATE: 08-OCT-1997
ATTORNEY/AGENT INFORMATION:
NAME: Doyle Leary Ph.D., Kathryn
REGISTRATION NUMBER: 36,317
REFERENCE/DOCKET NUMBER: 9596-1202
TELECOMMUNICATION INFORMATION:

; TELEPHONE: 215-965-1284
; TELEFAX: 215-567-2991
; TELEX: 831-494
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1621 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-972-927-3

Query Match 72.2%; Score 39; DB 3; Length 1621;
Best Local Similarity 66.7%; Pred. No. 8.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YMGGYW 7
DB 303 WVGFW 308

RESULT 11
US-08-972-927-6
; Sequence 6, Application US/08972927
; Patent No. 6168290
; GENERAL INFORMATION:
; APPLICANT: Rea, Philip A
; APPLICANT: Lu, Yu-Ping
; APPLICANT: Li, Ze-Sheng
; TITLE OF INVENTION: GLUTATHIONE-S-CONJUGATE TRANSPORT IN
; TITLE OF INVENTION: PLANTS
; NUMBER OF SEQUENCES: 20
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PANITCH SCHWARZE JACOBS & NADEL, P.C.
; STREET: One Commerce Square, 2005 Market Street, 22nd
; CITY: Philadelphia
; STATE: Pennsylvania
; COUNTRY: US
; ZIP: 19103-7086
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/972,927
; FILING DATE: 18-NOV-1997
; CLASSIFICATION: 800
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/031,040
; FILING DATE: 18-NOV-1996
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 60/061,328
; FILING DATE: 08-OCT-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Doyle Leary Ph.D., Kathryn
; REGISTRATION NUMBER: 36,317
; REFERENCE/DOCKET NUMBER: 9596-1202
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-965-1284
; TELEFAX: 215-567-2991
; TELEX: 831-494
; INFORMATION FOR SEQ ID NO: 6:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1622 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-972-927-6

Query Match 72.2%; Score 39; DB 3; Length 1622;

Best Local Similarity 66.7%; Pred. No. 8.1e+02;
Matches 4; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 2 YMGGYW 7
DB 303 WVGFW 308

RESULT 12
US-08-241-054-99
; Sequence 99, Application US/08241054
; Patent No. 5643873
; GENERAL INFORMATION:
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Cwirla, Steven E.
; APPLICANT: Dower, William J.
; APPLICANT: Koller, Kerry J.
; APPLICANT: Lee, Jung
; APPLICANT: Martens, Christine L.
; APPLICANT: Ruhland-Fritsch, Beatrice
; TITLE OF INVENTION: Peptides and Compounds That Bind
; TITLE OF INVENTION: Selectins Including Endothelium Leukocyte Adhesion
; TITLE OF INVENTION: Molecule 1
; NUMBER OF SEQUENCES: 169
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: 699 Prince Street
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: USA
; ZIP: 22313
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/241,054
; FILING DATE: 11-MAY-1994
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,295
; FILING DATE: 05-MAY-1993
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 07/881,395
; FILING DATE: 06-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Gerald F. Swiss
; REGISTRATION NUMBER: 30,113
; REFERENCE/DOCKET NUMBER: 000324-002
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-854-7400
; TELEFAX: 415-854-8275
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-241-054-99

Query Match 70.4%; Score 38; DB 1; Length 12;
Best Local Similarity 71.4%; Pred. No. 12;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYGGYW 7
DB 2 GVGGLW 8

RESULT 13
US-08-439-817-79

; Sequence 79, Application US/08439817
; Patent No. 5728802
; GENERAL INFORMATION:
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Cwirla, Steven E.
; APPLICANT: Dower, William J.
; APPLICANT: Koller, Kerry J.
; APPLICANT: Lee, Jung
; APPLICANT: Martens, Christine L.
; APPLICANT: Ruhland-Fritsch, Beatrice
; TITLE OF INVENTION: Peptides and Compounds That Bind
; TITLE OF INVENTION: Selectins Including Endothelial Leukocyte Adhesion
; TITLE OF INVENTION: Molecule I (ELAM-1)
; NUMBER OF SEQUENCES: 209
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Affymax Technologies, NV
; STREET: 4001 Miranda Ave.
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/439,817
; FILING DATE: 12-MAY-1995
; CLASSIFICATION: 530
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/241,054
; FILING DATE: 11-MAY-1994
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,295
; FILING DATE: 05-MAY-1993
; APPLICATION NUMBER: US 07/881,395
; FILING DATE: 06-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens, Lauren L.
; REGISTRATION NUMBER: 36,691
; REFERENCE/DOCKET NUMBER: 000324-046/1056.1
; TELEPHONE: 415-496-2300
; TELEFAX: 415-424-0832
; INFORMATION FOR SEQ ID NO: 79:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-439-817-79

Query Match 70.4%; Score 38; DB 1; Length 12;
Best Local Similarity 71.4%; Pred. No. 12;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYGGYV 7
Db 2 GVGGLW 8

RESULT 14
US-08-485-508-99
; Sequence 99, Application US/08485508
; Patent No. 5786322
; GENERAL INFORMATION:
; APPLICANT: Barrett, Ronald W.
; APPLICANT: Cwirla, Steven E.
; APPLICANT: Dower, William J.
; APPLICANT: Koller, Kerry J.

; APPLICANT: Lee, Jung
; APPLICANT: Martens, Christine L.
; APPLICANT: Ruhland-Fritsch, Beatrice
; TITLE OF INVENTION: Peptides and Compounds That Bind
; TITLE OF INVENTION: Selectins Including Endothelial Leukocyte Adhesion
; TITLE OF INVENTION: Molecule I
; NUMBER OF SEQUENCES: 162
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Affymax Technologies, NV
; STREET: 4001 Miranda Ave.
; CITY: Palo Alto
; STATE: California
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/485,508
; FILING DATE: Herewith
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/241,054
; FILING DATE: 11-MAY-1994
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US 08/057,295
; FILING DATE: 05-MAY-1993
; APPLICATION NUMBER: US 07/881,395
; FILING DATE: 06-MAY-1992
; ATTORNEY/AGENT INFORMATION:
; NAME: Stevens, Lauren L.
; REGISTRATION NUMBER: 36,691
; REFERENCE/DOCKET NUMBER: 000324-002/1056
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-496-2300
; TELEFAX: 415-424-0832
; INFORMATION FOR SEQ ID NO: 99:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 12 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
US-08-485-508-99

Query Match 70.4%; Score 38; DB 1; Length 12;
Best Local Similarity 71.4%; Pred. No. 12;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GYGGYV 7
Db 2 GVGGLW 8

RESULT 15
US-09-248-796A-22524
; Sequence 22524, Application US/09248796A
; Patent No. 6747137
; GENERAL INFORMATION:
; APPLICANT: Keith Weinstock et al
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.132
; CURRENT APPLICATION NUMBER: US/09/248,796A
; CURRENT FILING DATE: 1999-02-12
; PRIOR APPLICATION NUMBER: US 60/074,725
; PRIOR FILING DATE: 1998-02-13
; PRIOR APPLICATION NUMBER: US 60/096,409
; PRIOR FILING DATE: 1998-08-13

; NUMBER OF SEQ ID NOS: 28208
; SEQ ID NO 22524
; LENGTH: 86
; TYPE: PRT
; ORGANISM: Candida albicans
US-09-248-796A-22524

Query Match 70.4%; Score 38; DB 4; Length 86;
Best Local Similarity 66.7%; Pred. No. 73;
Matches 4; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YWGGYW 7
Db 10 FWGNYW 15

Search completed: April 18, 2005, 20:00:37
Job time : 28.5349 secs

True Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2005, 06:18:13 ; Search time 44 Seconds
(without alignments)
19.681 Million cell updates/sec

Title: SEQ2
Perfect score: 60
Sequence: 1 xgywgkgyw 9
Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues

Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79: *
1: pir1: *
2: pir2: *
3: pir3: *
4: pir4: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	80.0	1365	1 BVBVK5	killer toxin resis
2	45	75.0	441	2 F83694	short-chain fatty
3	44.5	74.2	83	2 E69803	hypothetical prote
4	44.5	74.2	187	2 G83047	hypothetical prote
5	44	73.3	358	2 H83554	hypothetical prote
6	43	71.7	80	2 AC2394	hypothetical prote
7	43	71.7	180	2 D83812	hypothetical prote
8	43	71.7	470	2 B72567	hypothetical prote
9	43	71.7	1311	2 A56390	mannosyl-glycopro
10	43	71.7	1312	2 E95006	beta-N-acetylhexos
11	43	71.7	1312	2 A97879	beta-N-acetylhexos
12	42	70.0	181	2 A83915	hypothetical prote
13	42	70.0	196	2 AE2918	acetyltransferase
14	42	70.0	207	2 H97692	conserved hypothet
15	42	70.0	375	2 AI2041	hypothetical prote
16	42	70.0	583	2 A69491	probable formylmet
17	42	70.0	844	2 I37079	DNA ligase (ATP) (
18	41	68.3	113	2 PH1663	Ig heavy chain v r
19	41	68.3	145	2 F75337	transposase - Dein
20	41	68.3	1447	2 S63669	UDPglucose-glycopr
21	40	66.7	102	2 B72742	hypothetical prote
22	40	66.7	194	2 H22845	hypothetical prote
23	40	66.7	228	2 S26009	hypothetical prote
24	40	66.7	250	1 S70854	probable heme tran
25	40	66.7	489	2 S77357	hypothetical prote
26	40	66.7	495	2 AD1927	hypothetical prote
27	40	66.7	534	1 A28716	sapoin precursor
28	40	66.7	589	2 A29476	muscarinic acetylch
29	40	66.7	589	2 B29514	muscarinic acetylch

30	40	66.7	590	2 S01114	muscarinic acetylch
31	40	66.7	590	2 S47572	muscarinic acetylch
32	40	66.7	655	1 ALKEG	cyclomaltodextrin
33	40	66.7	741	2 S73827	hypothetical prote
34	40	66.7	961	2 S67568	probable membrane
35	40	66.7	962	2 JC5808	G protein-coupled
36	39.5	65.8	273	2 E95268	probable ABC trans
37	39	65.0	115	2 B86328	protein F18014.29
38	39	65.0	177	2 D87288	acetyltransferase,
39	39	65.0	185	2 B97933	conserved hypothet
40	39	65.0	186	2 A95066	acetyltransferase,
41	39	65.0	247	2 T47883	secretory protein-
42	39	65.0	270	1 G83242	probable phosphoe
43	39	65.0	335	2 H84080	sugar ABC transpor
44	39	65.0	396	2 T04561	hypothetical prote
45	39	65.0	499	2 T36462	hypothetical prote

ALIGNMENTS

RESULT 1

BVBVK5

killer toxin resistance protein KRE5 precursor - yeast (Saccharomyces cerevisiae)
N/Alternate names: protein O6254; protein YOR336w

C/Species: Saccharomyces cerevisiae

C/Date: 30-Sep-1991 #sequence revision 10-May-1996 #text_change 09-Jul-2004

C/Accession: S62066; A36327; S67243; S71974; S12202

R/Parle, A.G.; Hand, N.J.; Goulding, S.G.; Wolfe, K.H.

submitted to the EMBL Data Library, June 1995

A/Description: Sequence of 29 kilobases around the PDR10 locus on the right arm of Sacch

A/Reference number: S62058

A/Accession: S62066

A/Molecule type: DNA

A/Residues: 1-1365 <PAB>

A/Cross-references: UNIPROT:P22023; EMBL:Z49821; NID:g1163062; PIDN:CAA89981.1; PID:g11

R/Headen, P.; Hill, K.; Wagner, J.; Slipetz, D.; Sommer, S.S.; Bussey, H.

Mol. Cell. Biol. 10, 3013-3019, 1990

A/Title: The yeast KRE5 gene encodes a probable endoplasmic reticulum protein required

A/Reference number: A36327; MUID:90258892; PMID:2188106

A/Accession: A36327

A/Molecule type: DNA

A/Residues: 1-581,583-779,'I',781,'IMMKQKQNIKSK',794,'K',795-1365 <MEA>

A/Cross-references: EMBL:M33556; NID:g171794; PIDN:AAA34725.1; PID:g171795

R/Goulding, S.E.; Hand, N.J.; Parle-McDermott, A.G.; Wolfe, K.H.

submitted to the Protein Sequence Database, July 1996

A/Reference number: S67233

A/Accession: S67243

A/Molecule type: DNA

A/Residues: 1-1365 <GOV>

A/Cross-references: EMBL:Z75244; NID:g1420730; PIDN:CAA99659.1; PID:g1420731; GSPDB:GNO

A/Experimental source: strain S288C

R/Parle-McDermott, A.G.; Hand, N.J.; Goulding, S.E.; Wolfe, K.H.

Yeast 12, 999-1004, 1996

A/Title: Sequence of 29 kb around the PDR10 locus on the right arm of Saccharomyces cer

A/Reference number: S71966; MUID:97051586; PMID:8896263

A/Accession: S71974

A/Status: nucleic acid sequence not shown; translation not shown

A/Molecule type: DNA

A/Residues: 1-1365 <PAW>

A/Cross-references: EMBL:Z49821; NID:g1163062; PIDN:CAA89981.1; PID:g1163070

A/Note: the nucleotide sequence was submitted to the EMBL Data Library, June 1995

C/Genetics:

A/Gene: SGD:KRE5; MIPS:YOR336w

A/Cross-references: SGD:S0005863; MIPS:YOR336w

A/Map position: 15R

C/Function:

A/Description: required for normal cell growth

A/Pathway: (1->6)-beta-D-glucan biosynthesis

C/Superfamily: KRE5 protein

C/Keywords: endoplasmic reticulum; glycoprotein

F;1-17/Domain: signal sequence #status predicted <SIG>

F;18-1365/Product: killer toxin resistance protein KRE5 #status predicted <MAT>

F:1362-1365/Region: endoplasmic reticulum retention signal #status predicted
F:115,228,293,457,519,523,604,644,870,1091,1150,1195/Binding site: carbohydrate (Asn) (C)

Query Match 80.0%; Score 48; DB 1; Length 1365;
Best Local Similarity 66.7%; Pred. No. 21;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWKGYYW 9
:|||||
Db 1233 DGYWKEGYW 1241

RESULT 2
F83694
short-chain fatty acids transporter atoE [imported] - Bacillus halodurans (strain C-125)
C:Species: Bacillus halodurans
C>Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
C:Accession: F83694
R:Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, P.; Hira
Nucleic Acids Res. 28, 4317-4331, 2000
A:Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
A:Reference number: A83650; MUID:20512582; PMID:11058132
A:Accession: F83694
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-441 <STO>
A:Cross-references: UNIPROT:Q9KFW3; GB:AP001508; GB:BA000004; NID:g10172890; PIDN:BA040
A:Experimental source: strain C-125
C:Genetics:
A:Gene: atoE
C:Superfamily: conserved hypothetical integral membrane protein HP0693

Query Match 75.0%; Score 45; DB 2; Length 441;
Best Local Similarity 71.4%; Pred. No. 20;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 YWKGYYW 9
:|||||
Db 48 YWGEYFW 54

RESULT 3
E69903
hypothetical protein yodI - Bacillus subtilis
C:Species: Bacillus subtilis
C>Date: 05-Dec-1997 #sequence_revision 05-Dec-1997 #text_change 09-Jul-2004
C:Accession: E69903
R:Kunst, F.; Ogasawara, N.; Moszer, I.; Albertini, A.M.; Alloni, G.; Azevedo, V.; Berter
C.; Bron, S.; Brouillet, S.; Brusch, C.V.; Caldwell, B.; Capuano, V.; Carter, N.M.; Chd
A.; Ehrlich, S.D.; Emmerson, P.T.; Entian, K.D.; Errington, J.; Fabbret, C.; Ferrari, E.
Nature 390, 249-256, 1997
A:Authors: Foulger, D.; Fritz, C.; Fujita, M.; Fujita, Y.; Fuma, S.; Galizzi, A.; Gallier
iech, J.; Harwood, C.R.; Henaut, A.; Hilbert, H.; Holsappel, S.; Hosono, S.; Hulio, M.F.
Koetter, P.; Konigstein, G.; Krogh, S.; Kumano, M.; Kurita, K.; Lapidus, A.; Lardinois,
A.; Muthers, J.; Ogiwara, A.; Oudega, B.; Lee, S.M.; Levine, A.; Liu, H.; Masuda, S.; Maueel
Y, M.; Ogawa, K.; Ogiwara, A.; Oudega, B.; Park, S.H.; Parro, V.; Pohl, T.M.; Portetelle
Rieger, M.; Rivolta, C.; Rocha, E.; Roche, B.; Rose, M.; Sadaie, Y.; Sato, T.; Scanlon,
A:Authors: Schleich, S.; Scoffone, R.; Scoffone, F.; Sekiguchi, J.; Sekowska, A.; Seron
akeuchi, M.; Tamakoshi, A.; Tanaka, T.; Terpsira, P.; Tognoni, A.; Tosato, V.; Uchiyama,
T.; Winters, P.; Wipat, A.; Yamamoto, H.; Yamane, K.; Yasumoto, K.; Yata, K.; Yoshida, K
A:Authors: Yoshikawa, H.F.; Zumanstern, E.; Yoshikawa, H.; Danchin, A.
A:Title: The complete genome sequence of the Gram-positive bacterium Bacillus subtilis.
A:Reference number: A69580; MUID:98044033; PMID:9384377
A:Accession: E69903
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-83 <KUN>
A:Cross-references: UNIPROT:O34654; GB:Z99114; GB:AL009126; NID:G2634230; PIDN:CAB13852.
A:Experimental source: strain 168
C:Genetics:
A:Gene: yodI

Query Match 74.2%; Score 44.5; DB 2; Length 83;

Best Local Similarity 77.8%; Pred. No. 4.8;
Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 XGYWKGYYW 9
:|||||
Db 52 YGYWG-GYW 59

RESULT 4
G83047
hypothetical protein PA4793 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: G83047
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: G83047
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-187 <STO>
A:Cross-references: UNIPROT:Q9HVL5; GB:AE004832; GB:AE004091; NID:g9951049; PIDN:AAG081,
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA4793

Query Match 74.2%; Score 44.5; DB 2; Length 187;
Best Local Similarity 77.8%; Pred. No. 10;
Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 XGYWKGYYW 9
:|||||
Db 108 GGYWG-GYW 115

RESULT 5
H83554
hypothetical protein PA0736 [imported] - Pseudomonas aeruginosa (strain PA01)
C:Species: Pseudomonas aeruginosa
C>Date: 15-Sep-2000 #sequence_revision 15-Sep-2000 #text_change 09-Jul-2004
C:Accession: H83554
R:Stover, C.K.; Pham, X.Q.; Erwin, A.L.; Mizoguchi, S.D.; Warrenner, P.; Hickey, M.J.; Br
adman, S.; Yuan, Y.; Brody, L.L.; Coulter, S.N.; Folger, K.R.; Kas, A.; Larbig, K.; Lim,
.; Lory, S.; Olson, M.V.
Nature 406, 959-964, 2000
A:Title: Complete genome sequence of Pseudomonas aeruginosa PA01, an opportunistic path
A:Reference number: A82950; MUID:20437337; PMID:10984043
A:Accession: H83554
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-358 <STO>
A:Cross-references: UNIPROT:Q9I5J3; GB:AE004508; GB:AE004091; NID:g9946611; PIDN:AAG041,
A:Experimental source: strain PA01
C:Genetics:
A:Gene: PA0736

Query Match 73.3%; Score 44; DB 2; Length 358;
Best Local Similarity 66.7%; Pred. No. 23;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWKGYYW 9
:|||||
Db 197 VGTWGGYW 205

RESULT 6
AC2394
hypothetical protein asr4707 [imported] - Nostoc sp. (strain PCC 7120)
C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120

C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C;Accession: AC2394
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriiguchi, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: AC2394
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-80 <KUR>
 A;Cross-references: UNIPROT:Q8YN63; GB:BA0000019; PIDN:BA076406.1; PID:g17133844; GSPDB:Q8YN63
 A;Experimental source: strain FCC 7120
 C;Genetics:
 A;Gene: asr4707

Query Match 71.7%; Score 43; DB 2; Length 80;
 Best Local Similarity 75.0%; Pred. No. 7.7; 1; Indels 0; Gaps 0;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWGKGY 8
 : |||||
 Db 37 GGYWGSY 44

RESULT 7
 D83812
 hypothetical protein BHI300 [imported] - Bacillus halodurans (strain C-125)
 C;Species: Bacillus halodurans
 C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 C;Accession: D83812
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hiraoka, Y.; Nakazawa, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. Nucleic Acids Res. 28, 4317-4331, 2000
 A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A;Reference number: A83650; MUID:20512582; PMID:11058132
 A;Accession: D83812
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-180 <STO>
 A;Cross-references: UNIPROT:Q9KDB3; GB:AP001511; GB:BA000004; NID:g10173727; PIDN:BA000004
 A;Experimental source: strain C-125
 C;Genetics:
 A;Gene: BHI300

Query Match 71.7%; Score 43; DB 2; Length 180;
 Best Local Similarity 75.0%; Pred. No. 17; 1; Indels 0; Gaps 0;
 Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWGKGY 8
 : |||||
 Db 100 QSYWGKGY 107

RESULT 8
 B72567
 hypothetical protein APE1820 - Aeropyrum pernix (strain K1)
 C;Species: Aeropyrum pernix
 C;Date: 20-Aug-1999 #sequence_revision 20-Aug-1999 #text_change 09-Jul-2004
 C;Accession: B72567
 R;Kawarabayashi, Y.; Hino, Y.; Horikawa, H.; Yamazaki, S.; Haikawa, Y.; Jin-no, K.; Takahawa, H.; Takamiya, M.; Masuda, S.; Funahashi, T.; Tanaka, T.; Kudoh, Y.; Yamazaki, J.; Kikuchi, Y.; Nakazawa, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S. DNA Res. 6, 83-101, 1999
 A;Title: Complete genome sequence of an aerobic hyper-thermophilic Crenarchaeon, Aeropyrum pernix
 A;Reference number: A72450; MUID:99310339; PMID:10382966
 A;Accession: B72567
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-470 <KAW>
 A;Cross-references: UNIPROT:Q9YAX5; DDBJ:AP0000062; NID:g5105244; PIDN:BA080823.1; PID:g5105244
 A;Experimental source: strain K1
 C;Genetics:
 A;Gene: APE1820
 C;Superfamily: phytoene dehydrogenase

Query Match 71.7%; Score 43; DB 2; Length 470;
 Best Local Similarity 83.3%; Pred. No. 41;
 Matches 5; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 4 WKGKGYW 9
 : |||||
 Db 329 WGRGYW 334

RESULT 9
 A56390
 mannosyl-glycoprotein endo-beta-N-acetylglucosaminidase (EC 3.2.1.96) precursor - Streptococcus pneumoniae
 C;Species: Streptococcus pneumoniae
 C;Date: 21-Jul-1995 #sequence_revision 21-Jul-1995 #text_change 09-Jul-2004
 C;Accession: A56390
 R;Clarke, V.A.; Platt, N.; Butters, T.D.
 J. Biol. Chem. 270, 8805-8814, 1995
 A;Title: Cloning and expression of the beta-N-acetylglucosaminidase gene from Streptococcus pneumoniae
 A;Reference number: A56390; MUID:95238375; PMID:7721787
 A;Accession: A56390
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1311 <CLA>
 A;Cross-references: UNIPROT:P49610; GB:L36923; NID:g784896; PIDN:AAC41450.1; PID:g784896
 C;Genetics:
 A;Gene: strH
 C;Keywords: glycosidase; hydrolase; tandem repeat
 F;1-33/Domain: signal sequence #status predicted <SIG>

Query Match 71.7%; Score 43; DB 2; Length 1311;
 Best Local Similarity 55.6%; Pred. No. 1.1e+02;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWGKGYW 9
 : |||||
 Db 868 ISYWSKGWW 876

RESULT 10
 E95006
 beta-N-acetylhexosaminidase [imported] - Streptococcus pneumoniae (strain TIGR4)
 C;Species: Streptococcus pneumoniae
 C;Date: 03-Aug-2001 #sequence_revision 03-Aug-2001 #text_change 09-Jul-2004
 C;Accession: E95006
 R;Tetelin, H.; Nelson, K.E.; Paulsen, I.T.; Eisen, J.A.; Read, T.D.; Peterson, S.; Heilbrunn, J.D.; Umayam, L.A.; White, O.; Salzberg, S.L.; Lewis, M.R.; Radune, D.; Holtzapple, M.; Hickey, E.K.; Holt, I.E.
 Science 293, 498-506, 2001
 A;Authors: Loftus, B.J.; Yang, F.; Smith, H.O.; Venter, J.C.; Dougherty, B.A.; Morrison, A.; Title: Complete Genome Sequence of a virulent isolate of Streptococcus pneumoniae.
 A;Reference number: A95000; MUID:21357209; PMID:11463916
 A;Accession: E95006
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1312 <KUR>
 A;Cross-references: UNIPROT:P49610; GB:AE005672; PIDN:AAK74246.1; PID:g14971522; GSPDB:AE005672
 A;Experimental source: strain TIGR4
 C;Genetics:
 A;Gene: SP0057

Query Match 71.7%; Score 43; DB 2; Length 1312;
 Best Local Similarity 55.6%; Pred. No. 1.1e+02;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWGKGYW 9
 : |||||
 Db 869 ISYWSKGWW 877

RESULT 11
 A97879
 beta-N-acetylhexosaminidase (EC 3.2.1.52) [imported] - Streptococcus pneumoniae (strain

C;Species: Streptococcus pneumoniae
 C;Date: 22-Oct-2001 #sequence_revision 22-Oct-2001 #text_change 09-Jul-2004
 C;Accession: A97879
 R;Hoskins, J.A.; Alborn Jr., W.; Arnold, J.; Blaszcak, L.; Burgett, S.; DeHoff, B.S.; E
 e, R.; LeBlanc, D.J.; Lee, L.N.; Lefkowitz, E.J.; Lu, J.; Matsushima, P.; McAhren, S.; M
 y, P.; Sun, P.M.; Winkler, M.E.
 J. Bacteriol. 183, 5709-5717, 2001
 A;Authors: Yang, Y.; Young-Bellido, M.; Zhao, G.; Zook, C.; Baltz, R.H.; Jaskunas, S.R.;
 A;Title: Genome of the Bacterium Streptococcus pneumoniae Strain R6.
 A;Reference number: A97872; MUID:21429245; PMID:11544234
 A;Accession: A97879
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-1312 <KUR>
 A;Cross-references: UNIPROT:Q8DRL6; GB:AE007317; PIDN:AAK98861.1; PID:G15457590; GSPDB:G
 C;Genetics:
 A;Gene: strH
 C;Keywords: glycosidase; hydrolase

Query Match 71.7%; Score 43; DB 2; Length 1312;
 Best Local Similarity 55.6%; Pred. No. 1.1e+02;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWKGYYW 9
 : |||||
 Db 869 ISYWSKGWW 877

RESULT 12
 A83915
 hypothetical protein BH2121 [imported] - Bacillus halodurans (strain C-125)
 C;Species: Bacillus halodurans
 C;Date: 01-Dec-2000 #sequence_revision 01-Dec-2000 #text_change 09-Jul-2004
 C;Accession: A83915
 R;Takami, H.; Nakasone, K.; Takaki, Y.; Maeno, G.; Sasaki, R.; Masui, N.; Fujii, F.; Hira
 Nucleic Acids Res. 28, 4317-4331, 2000
 A;Title: Complete genome sequence of the alkaliphilic bacterium Bacillus halodurans and
 A;Reference number: A83650; MUID:20512582; PMID:11058132
 A;Accession: A83915
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-181 <STO>
 A;Cross-references: UNIPROT:Q9KB15; GB:AP001514; GB:BA000004; NID:G10174613; PIDN:BA058
 A;Experimental source: strain C-125
 C;Genetics:
 A;Gene: BH2121
 C;Superfamily: ribosomal-protein-serine N-acetyltransferase

Query Match 70.0%; Score 42; DB 2; Length 181;
 Best Local Similarity 100.0%; Pred. No. 23;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 YWKGKY 8
 : |||||
 Db 100 YWKGKY 105

RESULT 13
 AE2918
 acetyltransferase [imported] - Agrobacterium tumefaciens (strain C58, Dupont)
 C;Species: Agrobacterium tumefaciens
 C;Date: 11-Jan-2002 #sequence_revision 11-Jan-2002 #text_change 09-Jul-2004
 C;Accession: AE2918
 R;Wood, D.W.; Serubal, J.C.; Kaul, R.; Monks, D.; Chen, L.; Wood, G.E.; Chen, Y.; Woo, I
 erage, G.; Gillet, W.; Grant, C.; Guenther, D.; Kutyavin, T.; Levy, R.; Li, M.; McClell
 i Karp, P.; Romero, P.; Zhang, S.
 Science 294, 2317-2323, 2001
 A;Authors: Yoo, H.; Tao, Y.; Biddle, P.; Jung, M.; Krespan, W.; Perry, M.; Gordon-Kamm,
 ster, E.W.
 A;Title: The Genome of the Natural Genetic Engineer Agrobacterium tumefaciens C58.
 A;Reference number: AB2577; MUID:21608550; PMID:11743193
 A;Accession: AE2918
 A;Status: preliminary

A;Molecule type: DNA
 A;Residues: 1-136 <KUR>
 A;Cross-references: UNIPROT:Q8UBR8; GB:AE008688; PIDN:AAAL43763.1; PID:G17741299; GSPDB:G
 A;Experimental source: strain C58 (Dupont)
 C;Genetics:
 A;Gene: Atu2782
 A;Map position: circular chromosome

Query Match 70.0%; Score 42; DB 2; Length 196;
 Best Local Similarity 100.0%; Pred. No. 25;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 YWKGKY 8
 : |||||
 Db 116 YWKGKY 121

RESULT 14
 H97692
 conserved hypothetical protein BH0837 (AP001510) [imported] - Agrobacterium tumefaciens
 C;Species: Agrobacterium tumefaciens
 C;Date: 30-Sep-2001 #sequence_revision 30-Sep-2001 #text_change 09-Jul-2004
 C;Accession: H97692
 R;Goodner, B.; Hinkle, G.; Gattung, S.; Miller, N.; Blanchard, M.; Ouello, B.; Goldman, B.;
 A.; Liu, F.; Wollam, C.; Allinger, M.; Doughty, D.; Scott, C.; Lappas, C.; Markelz, B.;
 Science 294, 2323-2328, 2001
 A;Title: Genome Sequence of the Plant Pathogen and Biotechnology Agent Agrobacterium tum
 A;Reference number: A97359; MUID:21608551; PMID:11743194
 A;Accession: H97692
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-207 <KUR>
 A;Cross-references: UNIPROT:Q8UBR8; GB:AE007869; PIDN:AAK88497.1; PID:G15158008; GSPDB:G
 C;Genetics:
 A;Gene: AGR_C_5049
 A;Map position: circular chromosome

Query Match 70.0%; Score 42; DB 2; Length 207;
 Best Local Similarity 100.0%; Pred. No. 26;
 Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 3 YWKGKY 8
 : |||||
 Db 127 YWKGKY 132

RESULT 15
 AI2041
 hypothetical protein all1887 [imported] - Nostoc sp. (strain PCC 7120)
 C;Species: Nostoc sp. PCC 7120
 A;Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
 C;Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
 C;Accession: AI2041
 R;Kaneko, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriguchi
 Nakazaki, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Tabata, T.
 DNA Res. 8, 205-213, 2001
 A;Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium An
 A;Reference number: AB1807; MUID:21595285; PMID:11759840
 A;Accession: AI2041
 A;Status: preliminary
 A;Molecule type: DNA
 A;Residues: 1-375 <KUR>
 A;Cross-references: UNIPROT:Q8YVT7; GB:BA000019; PIDN:BA073586.1; PID:G17130977; GSPDB:G
 A;Experimental source: strain PCC 7120
 C;Genetics:
 A;Gene: all1887

Query Match 70.0%; Score 42; DB 2; Length 375;
 Best Local Similarity 55.6%; Pred. No. 46;
 Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWKGYYW 9
 : |||||

Db 95 NGFWNGIW 103

Search completed: April 19, 2005, 06:43:17
Job time : 46 secs

Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2005, 05:59:25 ; Search time 176 Seconds
(without alignments)
26.186 Million cell updates/sec

Title: SEQ2
Perfect score: 60
Sequence: 1 xgywkggyw 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03.*

1: uniprot_sprot.*

2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	50	83.3	840	2 Q8A6N9	Q8A6N9 bacteroides
2	50	83.3	845	2 Q64QS3	Q64QS3 bacteroides
3	49	81.7	214	2 Q8A8U4	Q8A8U4 bacteroides
4	49	81.7	841	2 Q7MU18	Q7MU18 porphyromon
5	48	80.0	210	2 Q8A510	Q8A510 bacteroides
6	48	80.0	748	2 Q9HLG7	Q9HLG7 thermoplasma
7	48	80.0	1293	2 Q759B3	Q759B3 ashbya goss
8	48	80.0	1365	1 KRE5 YEAST	P22023 saccharomyc
9	46	76.7	574	1 ZBT3 HUMAN	Q9H5J0 homo sapien
10	45.5	75.8	74	2 Q63VF7	Q63VF7 burkholderi
11	45	75.0	198	2 Q9AXA8	Q9AXA8 oryza sativ
12	45	75.0	441	2 Q9KFW3	Q9KFW3 bacillus ha
13	45	75.0	1326	2 Q6FV20	Q6FV20 candida gla
14	44.5	74.2	83	1 YOD1 BACSU	Q34654 bacillus su
15	44.5	74.2	187	2 Q9HV15	Q9HV15 pseudomonas
16	44.5	74.2	189	2 Q8QB88	Q8QB88 pseudomonas
17	44.5	74.2	440	2 Q64VU6	Q64VU6 bacteroides
18	44	73.3	120	2 Q8BSY2	Q8BSY2 diachasmimo
19	44	73.3	358	2 Q915J3	Q915J3 pseudomonas
20	44	73.3	443	2 Q8EH39	Q8EH39 shewanella
21	44	73.3	716	2 Q9HLP3	Q9HLP3 thermoplasma
22	43.5	72.5	181	2 Q638P4	Q638P4 bacillus ce
23	43.5	72.5	182	2 Q734H0	Q734H0 bacillus ce
24	43	71.7	80	2 Q8YN63	Q8YN63 anabaena sp
25	43	71.7	127	2 Q858P1	Q858P1 lactococcus
26	43	71.7	133	2 Q21906	Q21906 bacterioph
27	43	71.7	180	2 Q9KDB3	Q9KDB3 bacillus ha
28	43	71.7	277	2 Q9F3T0	Q9F3T0 rhodothermu
29	43	71.7	314	2 Q8A188	Q8A188 bacteroides
30	43	71.7	446	2 Q7VX54	Q7VX54 bordetella
31	43	71.7	468	2 Q7WLG5	Q7WLG5 bordetella

32	43	71.7	470	2 Q9YAX5	Q9YAX5 aeropyrum p
33	43	71.7	479	2 Q7W819	Q7W819 bordetella
34	43	71.7	490	2 Q72R85	Q72R85 leptospira
35	43	71.7	490	2 Q8FAJ0	Q8FAJ0 leptospira
36	43	71.7	552	2 Q89JU4	Q89JU4 bradyrhizob
37	43	71.7	1312	1 STRH_STRPN	P49610 streptococc
38	43	71.7	1312	2 Q8DRL6	Q8DRL6 streptococc
39	42	70.0	126	2 Q9VNF1	Q9VNF1 drosophila
40	42	70.0	167	2 Q6HT29	Q6HT29 bacillus an
41	42	70.0	168	2 Q634H9	Q634H9 bacillus ce
42	42	70.0	168	2 Q730H0	Q730H0 bacillus ce
43	42	70.0	168	2 Q818A8	Q818A8 bacillus ce
44	42	70.0	168	2 Q81LM3	Q81LM3 bacillus an
45	42	70.0	168	2 Q6HDG0	Q6HDG0 bacillus th

ALIGNMENTS

```
RESULT 1
Q8A6N9
ID Q8A6N9 PRELIMINARY; PRT; 840 AA.
AC Q8A6N9;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DE Putative alanyl dipeptidyl peptidase.
GN OrderedLocusNames=BT1838;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12863928; DOI=10.1126/science.1080029;
RA Xu J., Bjursell M.K., Hmrod J., Deng S., Carmichael L.K.,
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL; A5016933; AAC76945.1; -
DR GO; GO:0003824; F: catalytic activity; IEA.
DR InterPro; IPR001005; Myb DNA binding.
DR InterPro; IPR001375; Peptidase S9.
DR InterPro; IPR006032; Ribosomal_S12_23.
DR InterPro; IPR000379; Ser_estr.
DR InterPro; IPR011042; TolB_C.
DR Pfam; PF00326; Peptidase_S9; 1.
DR PROSITE; PS00037; MYB_1; UNKNOWN 1.
DR PROSITE; PS00055; RIBOSOMAL_S12; UNKNOWN 1.
KW Complete proteome.
SQ SEQUENCE 840 AA; 94535 MW; 789B553158DE46A2/CRC64;
```

Query Match 83.3%; Score 50; DB 2; Length 840;
Best Local Similarity 66.7%; Pred. No. 31;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

Qy 1 XGYWKGYYW 9
Db 716 TSYWGEYYW 724

```
RESULT 2
Q64QS3
ID Q64QS3 PRELIMINARY; PRT; 845 AA.
AC Q64QS3;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DE Putative alanyl dipeptidyl peptidase.
GN ORFNames=BF3415;
OS Bacteroides fragilis.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
```

```

OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=817;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=YCH46;
RA Kuwahara T., Yamashita A., Hirakawa H., Nakayama H., Toh H., Okada N.,
RA Kuhara S., Hattori M., Hayashi T., Ohnishi Y.;
RT "Genomic analysis of Bacteroides fragilis reveals extensive DNA
RT inversions regulating cell surface adaptation.";
RL Proc. Natl. Acad. Sci. U.S.A. 0:0-0(2004).
DR EMBL; AP006841; BAD50158.1; -.
SQ SEQUENCE 845 AA; 95248 MW; 76A36642B8318D76 CRC64;

Query Match 83.3%; Score 50; DB 2; Length 845;
Best Local Similarity 66.7%; Pred. No. 32;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWGKGYW 9
Db : ||| |||
721 TSYWGEGYW 729

RESULT 3
Q8A8U4 PRELIMINARY; PRT; 214 AA.
AC Q8A8U4;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BT1073;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550859; PubMed=12663928; DOI=10.1126/science.1080029;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL; AE016930; AA076180.1; -.
KW Complete proteome.
SQ SEQUENCE 214 AA; 24735 MW; 84ABB75A740D226B CRC64;

Query Match 81.7%; Score 49; DB 2; Length 214;
Best Local Similarity 66.7%; Pred. No. 13;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWGKGYW 9
Db : ||| |||
124 NGWGFQYW 132

RESULT 4
Q7MU18 PRELIMINARY; PRT; 841 AA.
AC Q7MU18;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Conserved domain protein.
GN OrderedLocusNames=PG1754;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W83;
RX MEDLINE=22829867; PubMed=12949112;

```

```

RX DOI=10.1128/JB.185.18.5591-5601.2003;
RA Nelson K.E., Fleischmann R.D., Deboy R.T., Paulsen I.T., Fouts D.E.,
RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
RA Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
RT Porphyromonas gingivalis strain W83.";
RL J. Bacteriol. 185:5591-5601(2003).
DR EMBL; AE017178; AAO66756.1; -.
DR TIGR; PG1754; -.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000379; Ser_estrs.
DR Pfam; PF00326; Peptidase_S9; 1.
KW Complete proteome.
SQ SEQUENCE 841 AA; 94237 MW; A082DCD732EC3F36 CRC64;

Query Match 81.7%; Score 49; DB 2; Length 841;
Best Local Similarity 66.7%; Pred. No. 44;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWGKGYW 9
Db : ||| |||
717 SNYWGSGYW 725

RESULT 5
Q8A510 PRELIMINARY; PRT; 210 AA.
AC Q8A510;
DT 01-JUN-2003 (TrEMBLrel. 24, Created)
DT 01-JUN-2003 (TrEMBLrel. 24, Last sequence update)
DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Hypothetical protein.
GN OrderedLocusNames=BT2437;
OS Bacteroides thetaiotaomicron.
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Bacteroidaceae; Bacteroides.
OX NCBI_TaxID=818;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=VPI-5482 / ATCC 29148;
RX MEDLINE=22550858; PubMed=12663928; DOI=10.1126/science.1080029;
RA Xu J., Bjursell M.K., Himrod J., Deng S., Carmichael L.K.,
RA Chiang H.C., Hooper L.V., Gordon J.I.;
RT "A genomic view of the human-Bacteroides thetaiotaomicron symbiosis.";
RL Science 299:2074-2076(2003).
DR EMBL; AE016936; AAO77544.1; -.
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 210 AA; 23880 MW; 96392CE4D995FF2B CRC64;

Query Match 80.0%; Score 48; DB 2; Length 210;
Best Local Similarity 66.7%; Pred. No. 18;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWGKGYW 9
Db : ||| |||
124 PGYWGFSYW 132

RESULT 6
Q9HLG7 PRELIMINARY; PRT; 748 AA.
AC Q9HLG7;
DT 01-MAR-2001 (TrEMBLrel. 16, Created)
DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
DT 01-MAR-2001 (TrEMBLrel. 16, Last annotation update)
DE Hypothetical protein Ta0261.
GN OrderedLocusNames=Ta0261;
OS Thermoplasma acidophilum.
OC Archaea; Euryarchaeota; Thermoplasmatia; Thermoplasmatales;

```



```

OC Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=90258892; PubMed=2188106;
RA Meaden P., Hill K., Wagner J., Slipetz D., Sommer S.S., Bussey H.;
RT "The yeast KRES gene encodes a probable endoplasmic reticulum protein
RL required for (1--5)-beta-D-glucan synthesis and normal cell growth.";
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=97051586; PubMed=8896263;
RA Parle-Wdormort A.G., Hand N.J., Goulding S.G., Wolfe K.H.;
RT "Sequence of 29 kb around the PDR10 locus on the right arm of
RT Saccharomyces cerevisiae chromosome XV: similarity to part of
RT chromosome I.";
RL Yeast 12:999-1004 (1996).
CC -!- FUNCTION: Required for (1--5)-beta-D-glucan synthesis and normal
CC cell growth.
CC -!- SUBCELLULAR LOCATION: Endoplasmic reticulum lumen.
CC -!- SIMILARITY: Some, to D.melanogaster UGG.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC
CC -----
CC EMBL; M33556; AAA34725.1; -
CC EMBL; Z49821; CAA89981.1; -
CC EMBL; Z75244; CAA99659.1; -
CC PIR; S62066; BVBYK5.
CC GERMOnline; 143924; -.
CC SGD; S000005863; KRES.
CC GO; GO:0003980; F:UDP-glucose:glycoprotein glucosyltransferase. . . ; IDA.
CC PROSITE; PS000886; ER_TARGET_5.
CC INTERPRO; IPR000886; ER_TARGET_1.
CC KEGG; Cell wall; Endoplasmic reticulum; Glycoprotein; Signal.
KW SIGNAL
FT CHAIN 1 17 Potential.
FT CARBOHYD 18 1365 Killer toxin-resistance protein 5.
FT CARBOHYD 115 115 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 228 228 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 293 293 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 457 457 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 519 519 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 523 523 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 644 644 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 870 870 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1091 1091 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1150 1150 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 1195 1195 N-linked (GlcNAc...) (Potential).
FT SITE 1362 1365 Prevent secretion from ER.
FT CONFLICT 582 582 Missing (in Ref. 1).
FT CONFLICT 780 794 HUQNEVPTEHFEA -> ILTMKQKQKONISKAK (in Ref. 1).
SQ SEQUENCE 1365 AA; 156476 MW; D0F5851175CC0333 CRC64;

Query Match 80.0%; Score 48; DB 1; Length 1365;
Best Local Similarity 66.7%; Pred. No. 96;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWGKGYW 9
: ||| : |||
Db 1233 DGYWKEGYW 1241

RESULT 9
ZBT3 HUMAN STANDARD; PRT; 574 AA.
ID ZBT3 HUMAN
AC Q9H570;
DT 10-OCT-2003 (Rel. 42, Created)
DT 10-OCT-2003 (Rel. 42, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)

Thermoplasmataceae; Thermoplasma.
OX NCBI_TaxID=2303;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=DSM 1728;
RL MEDLINE=20479972; PubMed=11029001; DOI=10.1038/35035069;
RA Ruepp A., Graml W., Santos-Martinez M.-L., Koretke K.K., Volker C.,
RA Mewes H.-W., Frishman D., Stocker S., Lupas A.N., Baumeister W.;
RT "The genome sequence of the thermoacidophilic scavenger Thermoplasma
RT acidophilum.";
RL Nature 407:508-513 (2000).
DR EMBL; AL445063; CAC11406.1; -
KW Complete proteome; Hypothetical protein.
SQ SEQUENCE 748 AA; 82250 MW; 73A030A0E217EE39 CRC64;

Query Match 80.0%; Score 48; DB 2; Length 748;
Best Local Similarity 66.7%; Pred. No. 56;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWGKGYW 9
: ||| : |||
Db 530 SNTWKGYW 538

RESULT 7
Q759B3 PRELIMINARY; PRT; 1293 AA.
ID Q759B3;
AC Q759B3;
DT 05-JUL-2004 (T-EMBLrel. 27, Created)
DT 05-JUL-2004 (T-EMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (T-EMBLrel. 27, Last annotation update)
DE ADR3364WP.
GN ORFNames=ADR364W;
OS Ashbya gossypii (Yeast) (Eremothecium gossypii).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Eremothecium.
OX NCBI_TaxID=33169;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=ATCC 10895;
RA Voegeli S.E., Brachat S., Dietrich F.S., Lerch A., Gaffney T.,
RA Philippaen P.;
RL Submitted (SEP-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL; AE016817; AAS52284.1; -
DR AGD; ADR364W; -.
DR InterPro; IPR000977; DNA_ligase.
DR InterPro; IPR000886; ER_TARGET_5.
DR PROSITE; PS00697; DNA_LIGASE_A1; UNKNOWN_1.
DR PROSITE; PS00014; ER_TARGET; UNKNOWN_1.
SQ SEQUENCE 1293 AA; 145770 MW; 2FB177BDD5E93E37 CRC64;

Query Match 80.0%; Score 48; DB 2; Length 1293;
Best Local Similarity 66.7%; Pred. No. 92;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWGKGYW 9
: ||| : |||
Db 1172 HGYWEGYW 1180

RESULT 8
KRES YEAST STANDARD; PRT; 1365 AA.
ID KRES YEAST
AC P22023; Q12190;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Killer toxin-resistance protein 5 precursor.
GN Names=KRES; OrderedLocNames=YOR336W;
OS Saccharomyces cerevisiae (Baker's yeast).
OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
OC Saccharomycetales; Saccharomycetaceae; Saccharomyces.
OX NCBI_TaxID=4932;

```

DE Zinc finger and BTB domain containing protein 3.
GN Homo sapiens (Human)
OS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX NCBI_TaxID=9606;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Oca T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayaishi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Obayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoi T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,
RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Houta T.,
RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Inose N., Musashino K., Yuuki H., Oehima A., Sasaki N., Aotsuka S.,
RA Yohikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiya H., Satoh N., Takami S., Texashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakebe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujiwara T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs";
RL Nat. Genet. 36:40-45(2004).
RN [2]
RP SEQUENCE FROM N.A.
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner K.H., Shenmen C.F., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,
RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshiyuki S., Carninci P., Prange C.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
RA Bosak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahey J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Butterfield A.S.N., Krzywicki M.I., Skalska U., Smailus D.E.,
RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences";
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
CC -!- FUNCTION: May be involved in transcriptional regulation.
CC -!- SUBCELLULAR LOCATION: Nuclear (potential).
CC -!- SIMILARITY: Contains 1 BTB/POZ domain.
CC -!- SIMILARITY: Contains 2 C2H2-type zinc fingers.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way

CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
CC or send an email to license@isb-sib.ch).
CC -----
DR EMBL; AK027045; BAB15636.1; -;
DR EMBL; BC025249; AAH25249.1; -;
DR Genew; HGNC:22918; ZBTB3.
DR InterPro; IPR000210; BTB_POZ.
DR InterPro; IPR007087; Znf_C2H2.
DR Pfam; PF00051; BTB; 1.
DR Pfam; PF00096; zf-C2H2; 2.
DR SMART; SM00225; BTB; 1.
DR SMART; SM00355; Znf_C2H2; 2.
DR PROSITE; PS50097; BTB; 1.
DR PROSITE; PS00028; ZINC_FINGER_C2H2_1; 2.
DR PROSITE; PS0157; ZINC_FINGER_C2H2_2; 2.
KW DNA-binding; Metal-binding; Nuclear protein; Polymorphism; Repeat;
KW Transcription regulation; Zinc; Zinc-finger.
FT DOMAIN 74 142 BTB.
FT DOMAIN 227 346 Pro-rich.
FT ZN_FING 472 494 C2H2-type 1.
FT ZN_FING 500 523 C2H2-type 2.
FT VARIANT 574 574 I -> M (in dbSNP:544641).
FT /FTID=VAR_018382.
SQ SEQUENCE 574 AA; 61827 MW; 28C2FF4DB6C44036 CRC64;
Query Match 76.7%; Score 46; DB 1; Length 574;
Best Local Similarity 66.7%; Pred. No. 87;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;
QY 1 XGYWGKGYW 9
Db 213 SGHWGKGEW 221
RESULT 10
Q63VF7 PRELIMINARY; PRT; 74 AA.
ID Q63VF7 PRELIMINARY;
AC Q63VF7;
DT 25-OCT-2004 (TrEMBLrel. 28, Created)
DT 25-OCT-2004 (TrEMBLrel. 28, Last sequence update)
DT 25-OCT-2004 (TrEMBLrel. 28, Last annotation update)
DE Putative lipoprotein.
GN ORFNames=BPSL1288;
OS Burkholderia pseudomallei K96243.
OC Bacteria; Proteobacteria; Betaproteobacteria; Burkholderiales;
OC Burkholderiaceae; Burkholderia.
OX NCBI_TaxID=272560;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=K96243;
RX PubMed=15377794;
RA Holden M.T.G., Titball R.W., Peacock S.J., Cerdano-Tarraga A.M.,
RA Atkins T., Crossman L.C., Pitt T., Churcher C., Mungall K.,
RA Bentley S.D., Sebahia M., Thomson N.R., Bason N., Beacham I.R.,
RA Brooks K., Brown K.A., Brown N.F., Challis G.L., Cherevach I.,
RA Chillingworth T., Cronin A., Crosset B., Davis P., Deshaizer D.,
RA Feltham T., Fraser A., Hance Z., Hauser H., Holtroyd S., Jagels K.,
RA Keith K.E., Maddison M., Moule S., Price C., Quail M.A.,
RA Rabinowitch E., Rutherford K., Sanders M., Simmonds M.,
RA Songvilai S., Stevens K., Tumapa S., Vesaratchave M.,
RA Whitehead S., Yeats C., Barrell B.G., Oyston P.C.F., Parkhill J.;
RT "Genomic plasticity of the causative agent of melioidosis,
RT Burkholderia pseudomallei";
RL Proc. Natl. Acad. Sci. U.S.A. 101:14240-14245(2004).
DR EMBL; BX571965; CAH35282.1; -;
KW Lipoprotein.
SQ SEQUENCE 74 AA; 8055 MW; DC8B9521BDFDA72 CRC64;
Query Match 75.8%; Score 45.5; DB 2; Length 74;
Best Local Similarity 58.3%; Pred. No. 16;
Matches 7; Conservative 2; Mismatches 0; Indels 3; Gaps 1;

QY 1 XGTW---GKGYW 9
 Db :||| :|||
 60 FGWWSGGGGRGYW 71

RESULT 11

Q9AXA8 PRELIMINARY; PRT; 198 AA.
 AC Q9AXA8;
 DT 01-JUN-2001 (TrEMBLrel. 17, Created)
 DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
 DE P950IG01.9 protein.
 GN Name:P950IG01.9;
 OS Oryza sativa (Japonica cultivar-group).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
 OC Ehrhartoideae; Oryzeae; Oryza.
 OC NCBI_TaxID=39947;
 RN [1]

SEQUENCE FROM N.A.

RP RX PubMed=12447438; DOI=10.1038/nature01184;
 RA Sasaki T., Matsumoto T., Yamamoto K., Sakata K., Baba T., Katayose Y.,
 RA Wu J., Niimura Y., Cheng Z., Negamura Y., Antonio B.A., Kanamori H.,
 RA Hosokawa S., Masukawa M., Arikawa K., Chiden Y., Hayashi M.,
 RA Okamoto M., Ando T., Aoki H., Arita K., Hamada M., Harada C.,
 RA Hijishita S., Honda M., Ichikawa Y., Idonuma A., Iijima M., Ikeda M.,
 RA Ikeno M., Itoh S., Itoh Y., Itoh Y., Iwabuchi A., Kamiya K.,
 RA Karasawa W., Katagiri S., Kikuta A., Kobayashi N., Kono I.,
 RA Machita K., Maehara T., Mizuno H., Mizubayashi T., Mukai Y.,
 RA Nagasaki H., Nakashima M., Nakama Y., Nakamichi Y., Nakamura M.,
 RA Namiki N., Negishi M., Ohta I., Ono N., Saji S., Sakai K., Shibata M.,
 RA Shimokawa T., Shomura A., Song J., Takazaki Y., Terasawa K., Tsuji K.,
 RA Waki K., Yamagata H., Yamane H., Yoshiki S., Yoshihara R., Yukawa K.,
 RA Zhong H., Iwama H., Endo T., Ito H., Hahn J.H., Kim H.I., Eun M.Y.,
 RA Yano M., Jiang J., Gojobori T.;
 RT "The genome sequence and structure of rice chromosome 1.";
 RL Nature 420:312-316(2002).
 DR EMBL; AP002819; BAB21080.1; -;
 DR Gramene; Q9AXA8; -;

SEQUENCE 198 AA; 21475 MW; 13F337BABDB9BEF0 CRC64;

Query Match 75.0%; Score 45; DB 2; Length 198;
 Best Local Similarity 75.0%; Pred. No. 47;
 Matches 6; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 XGYWGKGY 8
 Db :||| :|||
 139 LGFWGKGY 146

RESULT 12

Q9KFW3 PRELIMINARY; PRT; 441 AA.
 AC Q9KFW3;
 DT 01-OCT-2000 (TrEMBLrel. 15, Created)
 DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
 DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
 DE Short-chain fatty acids transporter.
 GN Name:at0b; OrderedLocusNames=BH0358;
 OS Bacillus halodurans.
 OC Bacteria; Firmicutes; Bacilliales; Bacillaceae; Bacillus.
 OC NCBI_TaxID=86665;
 RN [1]

SEQUENCE FROM N.A.

RP STRAIN=C-125;
 RC MEDLINE=20512582; PubMed=11058132; DOI=10.1093/nar/28.21.4317;
 RA Takami H., Nakasone K., Takaki Y., Maeno G., Sasaki R., Masui N.,
 RA Fuji F., Hiramata C., Nakamura Y., Ogasawara N., Kuhara S.,
 RA Horikoshi K.;
 RT "Complete genome sequence of the alkaliphilic bacterium Bacillus
 RT halodurans and genomic sequence comparison with Bacillus subtilis.";
 RL Nucleic Acids Res. 28:4317-4331(2000).

DR EMBL; AP001508; BAB04077.1; -;
 DR FIR; F83694; F83694.
 DR GO; GO:0016020; C:membrane; IEA.
 DR GO; GO:0015635; F:short-chain fatty acid transporter activity; IEA.
 DR GO; GO:0015912; P:short-chain fatty acid transport; IEA.
 DR InterPro; IPR006160; SCFA_transpt.
 DR Pfam; PF02867; SCFA_trans; 1.
 DR ProDom; PD018634; SCFA_transpt; 1.
 KW Complete proteome.
 SQ SEQUENCE 441 AA; 47239 MW; 5387892F64302026 CRC64;

Query Match 75.0%; Score 45; DB 2; Length 441;
 Best Local Similarity 71.4%; Pred. No. 96;
 Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 3 YMGKGYW 9
 Db :||| :|||
 48 YWGEQFW 54

RESULT 13

Q6FV20 PRELIMINARY; PRT; 1326 AA.
 AC Q6FV20;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Candida glabrata strain CBS138 chromosome E complete sequence.
 GN ORFNames=CAGL0E054129;
 OS Candida glabrata CBS138.
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;
 OC Saccharomycetales; Mitosporic Saccharomycetales; Candida.
 OC NCBI_TaxID=284593;
 RN [1]

SEQUENCE FROM N.A.

RP STRAIN=CBS138;
 RC Genolevures;
 RA Dujon B., Sherman D., Fischer G., Durrens P., Casaregola S.,
 RA Lafontaine I., de Montigny J., Marck C., Neuvéglise C., Talla E.,
 RA Goffard N., Frangeul L., Aigle M., Anthouard V., Babour A., Barbe V.,
 RA Barnay S., Blanchin S., Beckerich J.M., Beyne E., Bleykasten C.,
 RA Boisrame A., Boyer J., Cattolico L., Confanieri F., de Daruvar A.,
 RA Despons L., Fabre E., Fairhead C., Ferry-Dumazet H., Groppi A.,
 RA Hantraye F., Hennequin C., Jauniaux N., Joyet P., Kachouri R.,
 RA Kerrest A., Koszul R., Lemaire M., Lesur I., Ma L., Muller H.,
 RA Nicaud J.M., Nikolski M., Oztas S., Ozier-Kalogeropoulos O.,
 RA Pellenz S., Potier S., Richard G.F., Straub M.L., Suleau A.,
 RA Swennene D., Tekalia F., Wesolowski-Louvel M., Westhof E., Wirth B.,
 RA Zeniou-Meyer M., Zivanovic I., Bolotin-Fukuhara M., Thierry A.,
 RA Bouchier C., Caudron B., Scarpelli C., Gaillardin C., Weissbach J.,
 RA Wincker P., Souciet J.L.;
 RT "Genome evolution in yeasts.";
 RL Nature 430:35-44(2004).
 DR EMBL; CR380951; CAG58843.1; -;
 SQ SEQUENCE 1326 AA; 152292 MW; 694BBE94391ED8DA CRC64;

Query Match 75.0%; Score 45; DB 2; Length 1326;
 Best Local Similarity 55.6%; Pred. No. 2.6e+02;
 Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWGKGYW 9
 Db :||| :|||
 1199 KGYWSEGHW 1207

RESULT 14

*YODI_BACSU
 ID YODI_BACSU STANDARD; PRT; 83 AA.
 AC O34654;
 DT 10-OCT-2003 (Rel. 42, Created)
 DT 10-OCT-2003 (Rel. 42, Last sequence update)
 DT 25-OCT-2004 (Rel. 45, Last annotation update)
 DE Hypothetical protein yodi.

GN Name=yodi; Synonyms=yolA; OrderedLocusNames=BSU19610;
 OS Bacillus subtilis.
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
 OX NCBI_TaxID=1423;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Wambutt R., Wedler H., Lapidus A., Sorokin A., Ehrlich S.D.;
 RT "Sequence analysis of the Bacillus subtilis chromosome region between
 RL the odhAB and spcC loci cloned in a yeast artificial chromosome.";
 RN Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Wambutt R., Jeong Y.-M., Choi S.-K., Park S.-H.;
 RT "Sequence analysis of the 30 kb region (182') of the Bacillus subtilis
 RL chromosome containing the cge cluster";
 RN Submitted (JUN-1997) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RC STRAIN=168;
 RX MEDLINE=98044033; PubMed=9384377; DOI=10.1038/36786;
 RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
 RA Arevalo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
 RA Borries R., Bruster L., Brans A., Braun M., Brignell S.C., Bron S.,
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
 RA Choi S.K., Codani J.J., Conneron I.F., Cummings N.J., Daniel R.A.,
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,
 RA Etian K.-D., Errington J., Fabret C., Ferrari E., Foulger D.,
 RA Fritz C., Fujita M., Fujita Y., Funa S., Gallizzi A., Galleron N.,
 RA Gim S.-Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,
 RA Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M.,
 RA Jones L.-M., Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M.,
 RA Klein C., Kobayashi Y., Koester P., Koningsstein G., Krogh S.,
 RA Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J.,
 RA Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Manuel C.,
 RA Medigue C., Medina N., Mellado R.P., Mizuno M., Mostl D., Nakai S.,
 RA Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B.,
 RA Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S.,
 RA Prescott A.M., Prasecan E., Pujic P., Purnelle B., Rapoport G.,
 RA Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R.,
 RA Scoffone F., Sekiguchi J., Sekowska A., Seror S.J., Serro P.,
 RA Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H.,
 RA Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P.,
 RA Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F.,
 RA Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H.,
 RA Weitzensegger T., Winters P., Wipat A., Yamamoto H., Yamane K.,
 RA Yasunoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E.,
 RA Yoshikawa H., Danchin A.;
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus
 RL subtilis.";
 RL Nature 390:249-256(1997).
 CC -----
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use by non-profit institutions as long as its content is in no way
 CC modified and this statement is not removed. Usage by and for commercial
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>
 CC or send an email to license@isb-sib.ch).
 CC -----
 CC EMBL; AF015775; AAB72056.1; -;
 DR EMBL; AF006665; AAB81166.1; -;
 DR EMBL; Z99114; CAB13852.1; -;
 DR PIR; E69903; E69903.
 DR Subtilist; BG13537; yodi.
 DR InterPro; IPR008991; Transl_SH3 like.
 KW Complete proteome; Hypothetical protein; Transmembrane.
 FT TRANSMEM 58 80 Potential.
 SQ SEQUENCE 83 AA; 9194 MW; 99F58EA2F0F36A43 CRC64;

Query Match

74.2%; Score 44.5; DB 1; Length 83;

Best Local Similarity 77.8%; Pred. No. 25;
 Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 XGYWKGWY 9
 :|||||
 DB 52 YGYWG-GYW 59

RESULT 15

Q9HV15 PRELIMINARY; PRT; 187 AA.
 ID Q9HV15;
 AC Q9HV15;
 DT 01-MAR-2001 (TrEMBLrel. 16, Created)
 DT 01-MAR-2001 (TrEMBLrel. 16, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 GN OrderedLocusNames=PA4793;
 OS Pseudomonas aeruginosa.
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;
 OC Pseudomonadaceae; Pseudomonas.
 OX NCBI_TaxID=287;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=ATCC 15692 / PA01;
 RX MEDLINE=20437337; PubMed=10984043; DOI=10.1038/35023079;
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,
 RA Garber R.L., Goltry L., Tolentino E., Westbrock-Wadman S., Yuan Y.,
 RA Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M.,
 RA Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T.,
 RA Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V.;
 RT "Complete genome sequence of Pseudomonas aeruginosa PA01, an
 RL opportunistic pathogen.";
 RL Nature 406:959-964(2000).
 DR EMBL; AE004892; AAG08179.1; -;
 DR PIR; G83047; G83047.
 KW Complete proteome; Hypothetical protein.
 SQ SEQUENCE 187 AA; 21281 MW; 89089EBEEC51897 CRC64;

Query Match 74.2%; Score 44.5; DB 2; Length 187;

Best Local Similarity 77.8%; Pred. No. 53;
 Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 XGYWKGWY 9
 :|||||
 DB 108 GGYWG-GYW 115

Search completed: April 19, 2005, 06:42:28

Job time : 179 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2005, 05:57:50 ; Search time 176 Seconds
(without alignments)
19.778 Million cell updates/sec

Title: SEQ2

Perfect score: 60

Sequence: 1 xgywkggyw 9

Scoring table: BLOSUM62DX
Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseqp1980s:*
- 2: Geneseqp1990s:*
- 3: Geneseqp2000s:*
- 4: Geneseqp2001s:*
- 5: Geneseqp2002s:*
- 6: Geneseqp2003as:*
- 7: Geneseqp2003bs:*
- 8: Geneseqp2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	60	100.0	9	3	Aay76816 Immunogen
2	59	98.3	8	2	Aay03714 Fluorine-
3	59	98.3	8	2	Aay03716 Fluorine-
4	59	98.3	8	7	Adg94002 Iodinated
5	59	98.3	8	8	Adl98013 Peptide h
6	51	85.0	8	3	Aay76818 Immunogen
7	49	81.7	841	3	Aab18511 H2 homolo
8	48	80.0	1365	2	Aar36780 KRE5. 7/1
9	48	80.0	1365	4	Aab72446 Yeast KRE
10	46	76.7	164	4	Aau16277 Human nov
11	46	76.7	164	6	Abus5346 Human nov
12	46	76.7	524	8	Adn99821 Novel hum
13	46	76.7	524	8	Adn99820 Novel hum
14	45	75.0	8	2	Aay03715 Fluorine-
15	45	75.0	8	3	Aay76817 Immunogen
16	45	75.0	8	7	Adg94005 Immunogen
17	45	75.0	8	8	Adl98014 Peptide h
18	45	75.0	152	4	Aab99111 Human pro
19	45	75.0	1090	4	Aam80087 Human pro
20	45	75.0	1180	4	Abg19170 Novel hum
21	44.5	74.2	339	7	Abu78095 Pseudomon
22	44.5	74.2	440	6	Abu20398 Pseudomon
23	44	73.3	383	7	Abu080711 Protein e
24	43	71.7	446	6	Abu23292 Protein e
25	43	71.7	960	8	Adp18467 S pneumoniae

26	43	71.7	1390	8	ADR95764	Adr95764 Novel S.
27	43	71.7	1303	8	ADK48522	Adk48522 Streptococ
28	43	71.7	1312	8	ADM92045	Adm92045 S pneumon
29	43	71.7	1329	3	AAY81779	Aay81779 Streptococ
30	43	71.7	1329	6	ABU02832	Abu02832 S. pneumo
31	42	70.0	71	5	ABF10992	Abf10992 Human ORF
32	42	70.0	121	4	AAU02601	Aau02601 Anti-adip
33	42	70.0	126	4	ABB69676	Abb69676 Drosophil
34	42	70.0	844	2	AAW69561	Aaw69561 Human lig
35	42	70.0	844	5	ABB77979	Abb77979 Amino aci
36	42	70.0	911	2	AAR95634	Aar95634 DNA-ligas
37	42	70.0	911	6	ABG72867	Abg72867 Human DNA
38	42	70.0	911	6	ABG75649	Abg75649 Human DNA
39	42	70.0	1891	4	ABG04331	Abg04331 Novel hum
40	41	68.3	50	8	ADQ11655	Adq11655 Myostatin
41	41	68.3	105	5	ABB77617	Abb77617 AnePV red
42	41	68.3	142	2	AAW09816	Aaw09816 PeLB-VH47
43	41	68.3	274	2	AAW09819	Aaw09819 VH4715-11
44	41	68.3	282	2	AAW09818	Aaw09818 VH4715-11
45	41	68.3	282	2	AAW35564	Aaw35564 HindIII-E

ALIGNMENTS

RESULT 1

AAAY76816

ID AAY76816 standard; peptide; 9 AA.

XX AAY76816;

XX 28-APR-2000 (first entry)

XX Immunogenic peptide for bi-specific antibody recognition.

XX Immunogenic peptide; bi-specific antibody; diagnosis; immune response;
XX diseased tissue identification; therapy.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "free amino acid group, protected amino acid

FT group, chelating agent or a metal-chelate complex"

FT Misc-difference 3 /note= "D-form residue"

FT Misc-difference 4 /note= "D-form residue"

FT Misc-difference 6 /note= "D-form residue"

FT /note= "D-form residue; modified with free amino acid

FT group, protected amino acid group, chelating agent or a

FT metal-chelate complex"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 9 /note= "D-form residue"

FT /note= "D-form residue"

WO9966951-A2.

29-DEC-1999.

22-JUN-1999; 99WO-US013879.

22-JUN-1998; 98US-0090142P.

14-OCT-1998; 98US-0104156P.

(IMMU-) IMMUNOMEDICS INC.

Hansen HJ, Griffiths GL, Leung S, McBride WJ, Qu Z;

WPI; 2000-160561/14.

Bi-specific antibodies that bind specific target tissue and targeted

XX 03-SEP-1998; 98WO-US018268.
 XX 03-SEP-1997; 97US-0057485P.
 XX (IMMU-) IMMUNOMEDICS INC.
 XX Griffiths GL;
 XX WPI; 1999-228967/19.
 XX Radiolabeling thiol-containing peptides with fluorine-18.
 XX Claim 15; Page 15; 22pp; English.
 XX The invention relates to a method for incorporating 18F radionuclide into
 CC peptide-containing targeting vectors for use in clinical positron
 CC emission tomography. Radiolabeling thiol-containing peptides with
 CC fluorine-18 (F-18) comprises reacting a peptide comprising a free thiol
 CC group with a labeling reagent of formula: 18F-(CH₂)_n-CRIR₂-(CH₂)_n-X, or a
 CC fluorinated alkene in which at least one of the two double bonded carbon
 CC atoms bears at least one leaving group comprising I, Br, Cl, azide,
 CC tosylate, mesylate, nosylate or triflate. n, m = 0-2; n+m = 0-2; X = I,
 CC Br, Cl, azide, tosylate, mesylate, nosylate, triflate, maleimide
 CC (optionally substituted by 1-2 alkyl) or 3-sulfonamide; R₁, R₂ = I,
 CC Br, Cl, azide, tosylate, mesylate, nosylate, triflate, H, CONH₂, COOH,
 CC OH, sulfonic acid, tertiary amine, quaternary ammonium, alkyl (optionally
 CC substituted by CONH₂, COOH, OH, sulfonic acid, tertiary amine or
 CC quaternary ammonium), COOR', CONR'2 or COR'; and R' = 1-6C alkyl or
 CC phenyl. The method is used for radiolabeling peptide-containing targeting
 CC vectors such as proteins, antibodies, antibody fragments and receptor-
 CC targeted peptides for use in routine clinical positron emission
 CC tomography. The method is simple and efficient. The method uses the
 CC unique property of the free thiol groups which are rapidly alkylated at
 CC neutral pH and moderate temperature. Sequences AY03714-716 represent
 CC examples of F-18 labeled peptides used in the method of detecting a
 CC tissue

XX SQ Sequence 8 AA;
 Query Match 98.3%; Score 59; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GYWGKGYW 9
 |||||
 Db 1 GYWGKGYW 8

RESULT 4
 ADG94002
 ID ADG94002 standard; peptide; 8 AA.
 AC ADG94002;
 XX 11-MAR-2004 (first entry)
 DT Iodinated immunogenic peptide.
 DE Iodinated immunogenic peptide; multi-specific antibody;
 KW polymer conjugate; tumour; cytostatic; photodynamic therapy.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 1
 FT /note= "Acetylated"
 FT Misc-difference 2
 FT /note= "D-form residue and optionally iodinated"
 FT Misc-difference 3
 FT /note= "D-form residue"
 FT Misc-difference 5
 FT /note= "Acetylated D-form residue"

FT Misc-difference 7 /note= "D-form residue and optionally iodinated"
 FT Misc-difference 8 /note= "D-form residue"
 XX US2003026764-A1.
 XX 06-FEB-2003.
 XX 31-JUL-2002; 2002US-00209592.
 XX 31-JUL-2001; 2001US-0308605P.
 XX (IMMU-) IMMUNOMEDICS INC.
 XX Griffiths GL;
 XX WPI; 2003-801085/75.
 XX Targeting an agent towards a target site in a tissue, by administering in
 FT a tissue, a multi-specific antibody or its fragment and a polymer
 FT conjugate that binds to the capture arm of the multi-specific antibody.
 XX Disclosure; Page 8; 19pp; English.
 XX The invention relates to targeting an agent towards a target site in a
 CC tissue, comprising administering to the tissue, a multi-specific antibody
 CC or its fragment, comprising a targeting arm that binds to an antigen of
 CC the target site and a capture arm that binds to a polymer conjugate, and
 CC administering a polymer conjugate that binds to the capture arm, the
 CC conjugate has a polymer conjugated to the agent such as therapeutic
 CC agent, a peptide, an enzyme and a labelled ligand. Also included is a kit
 CC useful for targeting a target site within a tissue in a subject or tissue
 CC sample comprising the above mentioned multi-specific antibody or its
 CC fragment and a polymer conjugate. The method is used for targeting an
 CC agent towards a target site in a tissue (e.g. a tumour). The method is
 CC also useful for therapeutic or diagnostic purposes and further in
 CC photodynamic therapy. The present sequence is an Iodinated immunogenic
 CC peptide used to raise antibodies which recognise only the iodinated
 CC peptide, used in the method of the invention.

XX SQ Sequence 8 AA;
 Query Match 98.3%; Score 59; DB 7; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GYWGKGYW 9
 |||||
 Db 1 GYWGKGYW 8

RESULT 5
 ADL98013
 ID ADL98013 standard; peptide; 8 AA.
 XX ADL98013;
 AC ADL98013;
 XX 20-MAY-2004 (first entry)
 DT Peptide hapten #1.
 DE
 XX
 XX photodynamic diagnosis; cancer; tumour; cardiovascular lesion;
 KW inflammatory disease; neurodegenerative disease; metabolic disease;
 KW infectious disease; B-cell malignancy; Alzheimer's disease; amyloidosis;
 KW autoimmune disease; bacterial infection; fungal infection;
 KW parasitic infection; viral infection;
 KW carcinoembryonic antigen-expressing tumour.
 XX Synthetic.
 OS
 XX Key Location/Qualifiers
 FH Modified-site 1

FT Modified-site 2 /note= "N-terminal acetyl"
 FT /note= "D form residue; optionally iodo Tyr"
 FT Misc-difference 3 /note= "D form residue"
 FT Modified-site 5 /note= "D form residue; acetylated"
 FT Modified-site 7 /note= "D form residue; optionally iodo Tyr"
 FT Misc-difference 8 /note= "D form residue"
 FT US2004043030-A1.
 PN 04-MAR-2004.
 PD
 PD
 PD 09-JUN-2003; 2003US-00455580.
 XX 31-JUL-2001; 2001US-0308605P.
 PR 31-JUL-2002; 2002US-00209592.
 XX (IMMU-) IMMUNOMEDICS INC.
 PA
 XX Griffiths GL, Goldenberg DM, Hansen HJ;
 XX WPI; 2004-313738/29.
 DR
 XX Treating cancer and metabolic diseases by administering a multi-specific
 PT antibody having a targeting arm that binds to an antigen and a capture
 PT arm that binds to a polymer conjugate comprising a therapeutic agent.
 XX
 PS Disclosure; Page 16; 24pp; English.
 XX
 CC The invention relates to a method of diagnosing or treating a disease or
 CC disorder. The method involves administering to a tissue a multi-specific
 CC antibody (I) or antibody fragment, comprising a targeting arm that binds
 CC to an antigen on the target site, and a capture arm that binds to a
 CC polymer conjugate, and administering to the tissue a polymer conjugate
 CC that binds to the capture arm, the polymer conjugate comprising a polymer
 CC conjugated to a diagnostic or therapeutic agent. Also included is a
 CC method for photodynamic diagnosis or treatment of a disease or disorder;
 CC or intravascular or endoscopic method for diagnosing or treating a
 CC disease or disorder. The method is useful for diagnosing or treating a
 CC disease or disorder chosen from cancer (oesophageal, gastric, colonic,
 CC rectal, pancreatic, lung, breast, ovarian, urinary bladder, endometrial,
 CC cervical, testicular, renal, adrenal and liver cancer, solid tumour, B-
 CC cell malignancy or T-cell malignancy); cardiovascular lesion; an
 CC inflammatory disease; neurodegenerative disease; metabolic disease; and
 CC an infectious disease. The B-cell malignancy is chosen from indolent
 CC forms of B-cell lymphomas, aggressive forms of B-cell lymphomas, chronic
 CC lymphatic leukaemias, acute lymphatic leukaemias, and multiple myeloma.
 CC The solid tumour is chosen melanoma, carcinoma (preferably renal
 CC carcinoma, lung carcinoma, intestinal carcinoma, and stomach carcinoma),
 CC glioma and sarcoma. The cardiovascular lesion is chosen from infarct,
 CC clot, embolus, atherosclerotic plaque and ischaemia. The
 CC neurodegenerative disease is Alzheimer's disease. The metabolic disease
 CC is amyloidosis, where the antibody binds amyloid. The disease or disorder
 CC is displaced or ectopic normal tissue chosen from endometrium, thymus,
 CC spleen and parathyroid. The method can be used for normal tissue
 CC ablation, where the tissue is chosen from bone marrow and spleen. The
 CC disease or disorder is an autoimmune disease such as myasthenia gravis,
 CC lupus nephritis, lupus erythematosus, and rheumatoid arthritis, Class III
 CC autoimmune diseases such as immune-mediated thrombocytopenias, such as
 CC acute idiopathic thrombocytopenic purpura and chronic idiopathic
 CC thrombocytopenic purpura, dermatomyositis, Sjogren's syndrome, multiple
 CC sclerosis, Sydenham's chorea, myasthenia gravis, systemic lupus
 CC erythematosus, lupus nephritis, rheumatic fever, polyglandular syndromes,
 CC bullous pemphigoid, diabetes mellitus, Henoch-Schönlein purpura, post-
 CC streptococcal nephritis, erythema nodosum, Takayasu's arteritis,
 CC Addison's disease, rheumatoid arthritis, sarcoidosis, ulcerative colitis,
 CC erythema multiforme, Iga nephropathy, polyarteritis nodosa, ankylosing
 CC spondylitis, Goodpasture's syndrome, thromboangitis obliterans, primary
 CC biliary cirrhosis, Hashimoto's thyroiditis, thyrotoxicosis, scleroderma,
 CC

CC chronic active hepatitis, polymyositis/dermatomyositis, polychondritis,
 CC pemphigus vulgaris, Wegener's granulomatosis, membranous nephropathy,
 CC amyotrophic lateral sclerosis, tabes dorsalis, giant cell
 CC arteritis/polyarthritis, pernicious anaemia, rapidly progressive
 CC glomerulonephritis, or fibrosing alveolitis. The infectious disease is
 CC chosen from bacterial, fungal, parasitic and viral lesion. The infectious
 CC disease is caused by a fungus chosen from Microsporum, Trichophyton,
 CC Epidermophyton, Sporothrix schenckii, Cryptococcus neoformans,
 CC Coccidioides immitis, Histoplasma capsulatum, Blastomyces dermatitidis,
 CC and Candida albicans. The infectious disease is caused by a virus chosen
 CC from HIV, herpes virus, cytomegalovirus, rabies virus, influenza virus,
 CC hepatitis B virus, Sendai virus, feline leukemia virus, Reo virus, polio
 CC virus, human serum parvo-like virus, simian virus 40, respiratory
 CC syncytial virus, mouse mammary tumour virus, Varicella-Zoster virus,
 CC Dengue virus, rubella virus, measles virus, adenovirus, human T-cell
 CC leukemia viruses, Epstein-Barr virus, murine leukemia virus, mumps virus,
 CC vesicular stomatitis virus, Sindbis virus, lymphocytic choriomeningitis
 CC virus, wart virus and blue tongue virus. The infectious disease is caused
 CC by a bacterium chosen from Bacillus anthracis, Streptococcus agalactiae,
 CC Legionella pneumophila, Streptococcus pyogenes, Escherichia coli,
 CC Neisseria gonorrhoeae, Neisseria meningitidis, Pneumococcus, Haemophilus
 CC influenzae B, Treponema pallidum, Lyme disease spirochetes, Pseudomonas
 CC aeruginosa, Mycobacterium leprae, Brucella abortus, Mycobacterium
 CC tuberculosis, and Tetanus toxin. The infectious disease is caused by a
 CC protozoa chosen from Plasmodium falciparum, Plasmodium vivax, Toxoplasma
 CC gondii, Trypanosoma rangeli, Trypanosoma cruzi, Trypanosoma
 CC rhodesiense, Trypanosoma brucei, Schistosoma mansoni, Schistosoma
 CC japonicum, Babesia bovis, Eimeria tenella, Onchocerca volvulus,
 CC Leishmania tropica, Trichinella spiralis, Onchocerca volvulus, Theileria
 CC parva, Taenia hydatigena, Taenia ovis, Taenia saginata, Echinococcus
 CC granulosus, and Mesocostoides corti. The infectious disease is caused by
 CC a mycoplasma chosen from Mycoplasma arthritis, M. hyorhinis, M. orale,
 CC M. arginini, Acholeplasma laidlawii, M. salivarium and M. pneumoniae. The
 CC cancer is preferably chosen from carcinoembryonic antigen (CEA) -
 CC expressing tumour or a CD20-expressing malignancy. The present sequence
 CC represents a peptide used in the method of the invention.
 XX
 SQ Sequence 8 AA;
 Query Match 98.3%; Score 59; DB 8; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 2 GYGKGYW 9
 DB 1 GYGKGYW 8
 RESULT 6
 AAY76818
 ID AAY76818 standard; peptide; 8 AA.
 XX
 AC AAY76818;
 XX
 DT 28-APR-2000 (first entry)
 XX
 DE Immunogenic peptide for bi-specific antibody recognition.
 XX
 KW Immunogenic peptide; bi-specific antibody; diagnosis; immune response;
 XX diseased tissue identification; therapy.
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Modified-site 1 /note= "acetylated"
 FT Misc-difference 2 /note= "D-form iodo-Tyr"
 FT Misc-difference 3 /note= "D-form residue"
 FT Misc-difference 5 /note= "D-form residue"
 FT Misc-difference 7 /note= "acetylated D-form residue"

FT Misc-difference 8 /note= "D-form iodo-Tyr"
 FT 7
 XX WO9966951-A2. /note= "D-form residue"
 XX 29-DEC-1999.
 XX 22-JUN-1999; 99WO-US013879.
 XX 22-JUN-1998; 98US-0090142P.
 PR 14-OCT-1998; 98US-0104156P.
 XX (IMMU-) IMMUNOMEDICS INC.
 PA Hansen HJ, Griffiths GL, Leung S, McBride WJ, Qu Z;
 PI WPI; 2000-160561/14.
 DR Bi-specific antibodies that bind specific target tissue and targeted
 XX conjugates.
 XX Claim 23; Page 61; 76pp; English.
 XX This sequence represents an immunogenic peptide for a bi-specific
 CC antibody. The invention relates to a method of treating or identifying
 CC diseased tissues in a patient comprising administering a bi-specific
 CC antibody (or fragment) having at least 1 arm (A) that specifically
 CC binds a targeted tissue and at least 1 arm (B) that specifically binds a
 CC targetable conjugate. The methods and bi-specific antibodies and fusion
 CC proteins are useful for pre-targeting methods of diagnosis and therapy.
 CC It is advantageous to raise bi-specific antibodies against a targetable
 CC conjugate that is capable of carrying at least 1 diagnostic or
 CC therapeutic agent. The characteristics of the chelator, metal chelate
 CC complex, therapeutic agent or diagnostic agent can be varied to
 CC accommodate differing applications without raising new bi-specific
 CC antibodies for each new application. The targetable conjugate is selected
 CC to elicit sufficient immune responses and also for rapid in vivo
 CC clearance when used within the bi-specific antibody targeting method
 XX
 XX Sequence 8 AA;
 SQ
 Query Match 85.0%; Score 51; DB 3; Length 8;
 Best Local Similarity 87.5%; Pred. No. 1.8e+06;
 Matches 7; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 2 GYWGKGYW 9
 DB |||||
 1 GYWHKGYW 8
 RESULT 7
 AAB18511
 ID AAB18511 standard; protein; 841 AA.
 AC AAB18511;
 XX 15-JAN-2001 (first entry)
 DT
 XX H2 homologue of prolyl-tripeptidyl peptidase DPP.
 DE
 XX Prolyl tripeptidyl-peptidase; amidolytic activity; periodontal disease;
 KW gingivitis; periodontitis.
 KW Porphyromonas gingivalis.
 XX
 OS WO200052147-A2.
 XX
 XX 08-SEP-2000.
 PD
 XX 03-MAR-2000; 2000WO-US005551.
 PF
 XX 05-MAR-1999; 99US-0123148P.
 PR

XX (UYGE-) UNIV GEORGIA RES FOUND INC.
 PA (TRAV/) TRAVIS J.
 PA (POTE/) POTEPA J.
 PA (BANE/) BANBULA A.
 XX Travis J, Potempa J, Banbula A;
 PI WPI; 2000-594181/56.
 XX Prolyl tripeptidyl-peptidase, active analog, fragment or variant useful
 PT for identifying its inhibitor which is useful for protecting an animal
 PT from a periodontal disease such as gingivitis and periodontitis.
 XX
 XX Claim 22; Fig 6; 58pp; English.
 PS
 XX The present sequence represents a H2 homologue of a prolyl tripeptidyl-
 CC peptidase (designated DPP) from Porphyromonas gingivalis. The prolyl
 CC tripeptidyl-peptidase has an amidolytic activity, and cleaves a peptide
 CC bond in a target polypeptide having at least 4 amino acids. This bond is
 CC between a proline and an amino acid attached to the alpha-carboxyl group
 CC end of the proline. The polypeptide is useful for identifying inhibitors.
 CC These inhibitors are then useful for reducing the growth of bacterium or
 CC for protecting an animal from a periodontal disease such as gingivitis
 CC and periodontitis caused by Porphyromonas gingivalis
 XX
 XX Sequence 841 AA;
 SQ
 Query Match 81.7%; Score 49; DB 3; Length 841;
 Best Local Similarity 66.7%; Pred. No. 1.4e+02;
 Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;
 QY 1 XGVGKGYW 9
 DB :|||
 717 SNVWGSYW 725
 RESULT 8
 AAR36780
 ID AAR36780 standard; protein; 1365 AA.
 XX AAR36780;
 AC
 XX 16-JUL-1993 (first entry)
 DT
 XX KRE5.
 DE
 XX Yeast; cell wall; beta-glucan; assembly; pathway; KRE1; KRE5; growth;
 KW secretory; O-linked mannose; (1-6)-beta-glucan; epistasis; morphology;
 KW hydrophilic; glycoprotein; COOH-terminal; endoplasmic reticulum; ER;
 KW retention signal; antifungal agent.
 XX
 OS Saccharomyces cerevisiae.
 OS
 XX US5194600-A.
 PN
 XX 16-MAR-1993.
 PD
 XX 05-MAR-1990; 90US-00488316.
 PF
 XX 05-MAR-1990; 90US-00488316.
 PR
 XX (ROYA-) ROYAL INST ADVANCEMENT LEARNING.
 PA
 XX Bussey H, Boone C, Sommer SS, Hill K, Meaden P;
 PI WPI; 1993-109384/13.
 XX N-PSDB; AAQ38899.
 DR
 XX New DNA encoding genes which participate in beta-glucan assembly - useful
 PT for producing mutants for in-vivo screening of antifungal agents and
 PT providing tools for in-vitro screening.
 XX

PS Claim 1; Col 38-44; 24pp; English.

XX The sequences given in AAR34785 and AAR36780 represent proteins which

CC participate in a yeast cell wall beta-glucan assembly pathway. These

CC proteins represent KRE1 and KRE5 respectively, and are essential for

CC normal cell growth. KRI is a Ser/Thr rich protein that is directed into

CC the yeast secretory pathway, where it is highly modified, probably

CC through addition of O-linked mannose residues. Gene disruption of the

CC KRE1 locus leads to a 40% reduced level of cell wall (1-6)-beta-glucan.

CC Mutations at KRE5 also caused defects in cell wall (1-6)-beta-glucan

CC production and appears to be epistatic to KRE1. KRE5 is a large

CC hydrophilic secretory glyco- protein which contains the COOH-terminal

CC endoplasmic reticulum (ER) retention signal (His-Asp-Glu-Leu). Deletion

CC of the KRE5 gene results in cells with aberrant morphology and extremely

CC compromised growth. KRE1 and KRE5 are useful as tools for the in vitro

CC screening of anti- fungal agents which inhibit fungi pathogenic to plants

CC and animals. The genes can be used to produce mutants for in vivo

CC screening of antifungal agents

XX SQ Sequence 1365 AA;

Query Match 80.0%; Score 48; DB 2; Length 1365;

Best Local Similarity .66.7%; Pred. No. 3.1e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWKGKGYW 9

Db 1233 DGYWKEGYW 1241

RESULT 9

AAB72446

ID AAB72446 standard; protein; 1365 AA.

XX AC AAB72446;

XX DT 08-MAY-2001 (first entry)

XX DE Yeast KRE5.

XX UGGT; UDP-glucose:glycoprotein glucosyltransferase; enzyme; KRE5.

XX OS Saccharomyces cerevisiae.

XX PN WO200112845-A1.

XX PD 22-FEB-2001.

XX PF 27-JUL-2000; 2000WO-CA000883.

XX PR 18-AUG-1999; 99US-00376330.

XX PA (CANADA) NAT RES COUNCIL CANADA.

XX PI Tessier DC, Dignard D, Bergeron JMM, Thomas DY;

XX DR WPI; 2001-218358/22.

XX Determining the effect of a test sample on UDP-glucose:glycoprotein

FT glucosyltransferase (UGGT), useful for measuring UGGT activity, comprises

PT exposing an acceptor substrate for UGGT to a labeled donor in the

PT presence of UGGT.

XX Disclosure; Fig 9; 95pp; English.

XX The present invention relates to a method for determining the effect of a

CC test sample on UDP-glucose:glycoprotein glucosyltransferase (UGGT)

CC activity. The method comprises exposing an acceptor substrate for UGGT to

CC a labelled donor in the presence of the test sample and UGGT. The method

CC is useful for determining UGGT activity. In particular, the method is

CC useful in glucosyltransferase assay and kinetics measurement for

CC determining UGGT activity. UGGT is a soluble enzyme of the endoplasmic

CC reticulum which catalyses the addition of a glucose residue onto

CC asparagine-linked oligosaccharides, which are present on incorrectly

CC folded glycoproteins. The present sequence was used in a sequence

CC homology comparison with rat UGGT (see AAR60732 and AAB72436) which was

CC used in the method of the present invention

XX SQ Sequence 1365 AA;

Query Match 80.0%; Score 48; DB 4; Length 1365;

Best Local Similarity .66.7%; Pred. No. 3.1e+02;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWKGKGYW 9

Db 1233 DGYWKEGYW 1241

RESULT 10

AAU16277

ID AAU16277 standard; protein; 164 AA.

XX AC AAU16277;

XX DT 07-NOV-2001 (first entry)

XX DE Human novel secreted protein, Seq ID 1230.

XX Human; immunosuppressive; antiarthritic; antirheumatic; cytostatic;

KW cardiant; vasotropic; cerebroprotective; nootropic; neuroprotective;

KW antibacterial; virucide; fungicide; ophthalmological; vulnerary;

KW secreted protein; rheumatoid arthritis; hyperproliferative disorder;

KW cardiovascular disorder; cardiac arrest; cerebrovascular disorder;

KW cerebral ischaemia; angiogenesis; nervous system disorder;

KW Alzheimer's disease; infection; ocular disorder; corneal infection;

KW wound healing; epithelial cell proliferation; skin ageing; food additive;

KW preservative; antiproliferative.

XX OS Homo sapiens.

XX PN WO200155322-A2.

XX PD 02-AUG-2001.

XX PF 17-JAN-2001; 2001WO-US001341.

XX PR 31-JAN-2000; 2000US-0179065P.

XX PR 04-FEB-2000; 2000US-0180628P.

XX PR 24-FEB-2000; 2000US-0184664P.

XX PR 02-MAR-2000; 2000US-0186350P.

XX PR 16-MAR-2000; 2000US-0189874P.

XX PR 17-MAR-2000; 2000US-0190076P.

XX PR 18-APR-2000; 2000US-0198123P.

XX PR 19-MAY-2000; 2000US-0205515P.

XX PR 07-JUN-2000; 2000US-0209467P.

XX PR 28-JUN-2000; 2000US-0214886P.

XX PR 30-JUN-2000; 2000US-0215135P.

XX PR 07-JUL-2000; 2000US-0216647P.

XX PR 07-JUL-2000; 2000US-0218880P.

XX PR 11-JUL-2000; 2000US-0217487P.

XX PR 11-JUL-2000; 2000US-0217496P.

XX PR 14-JUL-2000; 2000US-0218290P.

XX PR 26-JUL-2000; 2000US-0220963P.

XX PR 26-JUL-2000; 2000US-0220964P.

XX PR 14-AUG-2000; 2000US-0224518P.

XX PR 14-AUG-2000; 2000US-0224519P.

XX PR 14-AUG-2000; 2000US-0225213P.

XX PR 14-AUG-2000; 2000US-0225214P.

XX PR 14-AUG-2000; 2000US-0225266P.

XX PR 14-AUG-2000; 2000US-0225267P.

XX PR 14-AUG-2000; 2000US-0225268P.

XX PR 14-AUG-2000; 2000US-0225270P.

XX PR 14-AUG-2000; 2000US-0225447P.

XX PR 14-AUG-2000; 2000US-0225757P.

XX PR 14-AUG-2000; 2000US-0225758P.

PR 14-AUG-2000; 2000US-0225759P.
PR 18-AUG-2000; 2000US-0226279P.
PR 22-AUG-2000; 2000US-0226681P.
PR 22-AUG-2000; 2000US-0226868P.
PR 22-AUG-2000; 2000US-0227182P.
PR 23-AUG-2000; 2000US-0227009P.
PR 30-AUG-2000; 2000US-0228924P.
PR 01-SEP-2000; 2000US-0229287P.
PR 01-SEP-2000; 2000US-0229343P.
PR 01-SEP-2000; 2000US-0229344P.
PR 01-SEP-2000; 2000US-0229345P.
PR 05-SEP-2000; 2000US-0229509P.
PR 05-SEP-2000; 2000US-0229513P.
PR 06-SEP-2000; 2000US-0230437P.
PR 08-SEP-2000; 2000US-0231242P.
PR 08-SEP-2000; 2000US-0231243P.
PR 08-SEP-2000; 2000US-0231244P.
PR 08-SEP-2000; 2000US-0231413P.
PR 08-SEP-2000; 2000US-0231414P.
PR 08-SEP-2000; 2000US-0232080P.
PR 08-SEP-2000; 2000US-0232081P.
PR 12-SEP-2000; 2000US-0231968P.
PR 14-SEP-2000; 2000US-0232337P.
PR 14-SEP-2000; 2000US-0232338P.
PR 14-SEP-2000; 2000US-0232339P.
PR 14-SEP-2000; 2000US-0232400P.
PR 14-SEP-2000; 2000US-0232401P.
PR 14-SEP-2000; 2000US-0233063P.
PR 14-SEP-2000; 2000US-0233064P.
PR 14-SEP-2000; 2000US-0233065P.
PR 21-SEP-2000; 2000US-0234223P.
PR 21-SEP-2000; 2000US-0234274P.
PR 25-SEP-2000; 2000US-0234997P.
PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 08-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.

PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249264P.
PR 17-NOV-2000; 2000US-0249265P.
PR 17-NOV-2000; 2000US-0249297P.
PR 17-NOV-2000; 2000US-0249299P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 05-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.
XX

PI Rosen CA, Barash SC, Ruben SM;

XX WPI; 2001-488783/53.
XX N-PSDB; AAS26284.

PT New nucleic acid molecules encoding 461 human secreted proteins for
diagnosing, preventing, treating or ameliorating medical conditions and
used as food additives or preservatives.

XX Claim 11; SEQ ID NO 1230; 980pp; English.

XX The invention relates to isolated nucleic acid molecules and their
encoded secreted proteins. The nucleic acids and proteins are used to
prevent, treat or ameliorate a medical condition in e.g. humans, mice,
rabbits, goats, horses, cats, dogs, chickens or sheep. They are also used
in diagnosing a pathological condition or susceptibility to a
pathological condition. Antibodies to the proteins can also be used in
alleviating symptoms associated with the disorders and in diagnostic
immunoassays e.g. radioimmunoassays or enzyme linked immunosorbent assays
(ELISA). Disorders which are diagnosed or treated include autoimmune
diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g.
neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac
arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis,
nervous system disorders e.g. Alzheimer's disease, infections caused by
bacteria, viruses and fungi and ocular disorders e.g. corneal infection,
and many other disorders listed in the specification. The polypeptides
can also be used to aid wound healing and epithelial cell proliferation,
to prevent skin aging due to sunburn, to maintain organs before
transplantation, for supporting cell culture of primary tissues, to
regenerate tissues and in chemotaxis. The polypeptides can also be used
as a food additive or preservative to increase or decrease storage
capabilities, fat content, lipid, protein, carbohydrate, vitamins,
minerals, cofactors and other nutritional components. The present
sequence represents a novel secreted protein of the invention. Note: The
sequence data for this patent did not form part of the printed

```

Query Match      76.7%; Score 46; DB 4; Length 164;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 XGYWGKGYW 9
Db      124 SCHWKGW 132

RESULT 11
ABU55346
ID      ABU55346 standard; protein; 164 AA.
XX
XX      ABU55346;
AC
XX      DT 18-MAR-2003 (first entry)
XX
XX      DE Human novel polypeptide #433.
XX
XX      Human; neural disorder; immune system disorder; renal disorder;
KW      muscular disorder; respiratory disease; reproductive disorder;
KW      gastrointestinal disorder; pulmonary disorder; cardiovascular disorder;
KW      hyperproliferative disorder; inflammatory disease; allergic reaction;
KW      blood related disorder; cancer; immunosuppressive; antinflammatory;
KW      cardiovascular; nephrotropic; cytostatic; antiallergic; thrombolytic;
KW      haemostatic; antiarteriosclerotic.
XX
XX      OS Homo sapiens.
XX
XX      US2002132753-A1.
XX
XX      PD 19-SEP-2002.
XX
XX      PF 17-JAN-2001; 2001US-00764864.
XX
XX      PR 31-JAN-2000; 2000US-0179065P.
XX      PR 04-FEB-2000; 2000US-0180628P.
XX      PR 28-JUN-2000; 2000US-0214986P.
XX      PR 07-JUL-2000; 2000US-0216647P.
XX      PR 07-JUL-2000; 2000US-0216980P.
XX      PR 11-JUL-2000; 2000US-0217487P.
XX      PR 11-JUL-2000; 2000US-0217496P.
XX      PR 14-JUL-2000; 2000US-0218290P.
XX      PR 26-JUL-2000; 2000US-0220963P.
XX      PR 26-JUL-2000; 2000US-0220964P.
XX      PR 14-AUG-2000; 2000US-0224518P.
XX      PR 14-AUG-2000; 2000US-0224519P.
XX      PR 14-AUG-2000; 2000US-0225267P.
XX      PR 14-AUG-2000; 2000US-0225268P.
XX      PR 14-AUG-2000; 2000US-0225270P.
XX      PR 14-AUG-2000; 2000US-0225447P.
XX      PR 14-AUG-2000; 2000US-0225757P.
XX      PR 14-AUG-2000; 2000US-0225758P.
XX      PR 22-AUG-2000; 2000US-0226868P.
XX      PR 30-AUG-2000; 2000US-0228924P.
XX      PR 01-SEP-2000; 2000US-0229287P.
XX      PR 01-SEP-2000; 2000US-0229343P.
XX      PR 01-SEP-2000; 2000US-0229344P.
XX      PR 01-SEP-2000; 2000US-0229345P.
XX      PR 05-SEP-2000; 2000US-0229509P.
XX      PR 05-SEP-2000; 2000US-0229513P.
XX      PR 08-SEP-2000; 2000US-0231413P.
XX      PR 21-SEP-2000; 2000US-0234232P.
XX      PR 21-SEP-2000; 2000US-0234274P.
XX      PR 25-SEP-2000; 2000US-0234597P.
XX      PR 27-SEP-2000; 2000US-0235834P.
XX      PR 29-SEP-2000; 2000US-0236327P.
XX      PR 29-SEP-2000; 2000US-0236367P.
XX      PR 29-SEP-2000; 2000US-0236368P.
XX      PR 29-SEP-2000; 2000US-0236369P.
XX      PR 29-SEP-2000; 2000US-0236370P.
XX      PR 02-OCT-2000; 2000US-0236802P.
XX      PR 02-OCT-2000; 2000US-0237037P.

PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 02-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239353P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241809P.
PR 01-NOV-2000; 2000US-0244617P.
PR 17-NOV-2000; 2000US-0249299P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
XX
XX      (ROSE/) ROSEN C A.
PA      (RUBE/) RUBEN S M.
PA      (BARA/) BARASH S C.
XX
XX      PI Rosen CA, Ruben SM, Barash SC;
XX
XX      WPI; 2003-147444/14.
DR      N-PSDB; ABX73605.
XX
XX      New polypeptides and nucleic acids, useful in gene therapy for treating,
PT      inhibiting or preventing e.g. neural, immune system, muscular,
PT      respiratory, reproductive, gastrointestinal, pulmonary, cardiovascular or
PT      renal disorders.
XX
XX      Claim 11; SEQ ID NO 1230; 402pp; English.
PS
XX
XX      The invention relates to human novel polypeptides and their associated
CC      polynucleotides. The polypeptides and polynucleotides are useful in gene
CC      therapy for treating, inhibiting or preventing neural disorders, immune
CC      system disorders (e.g. systemic lupus erythematosus, rheumatoid arthritis
CC      and multiple sclerosis), muscular disorders, respiratory diseases (e.g.
CC      nasal vestibulitis, nasal polyps and sinusitis), reproductive disorders,
CC      gastrointestinal disorders, pulmonary disorders, cardiovascular disorders,
CC      congenital heart defects, Ebstein's anomaly and hypoplastic left
CC      heart syndrome), renal disorders (e.g. acute kidney failure and end-stage
CC      renal disease), hyperproliferative disorders (e.g. Hodgkin's disease and
CC      leukaemia), inflammatory diseases (e.g. septic shock, bursitis and
CC      appendicitis), allergic reactions and conditions (e.g. asthma), blood
CC      related disorders (e.g. thrombosis, atherosclerosis and myocardial
CC      infarction) and cancerous diseases. Sequences ABU54914-ABU55699 and
CC      ABU55748 represent human novel polypeptides of the invention
XX
XX      SQ Sequence 164 AA;

Query Match      76.7%; Score 46; DB 6; Length 164;
Best Local Similarity 66.7%; Pred. No. 78;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY      1 XGYWGKGYW 9
Db      124 SCHWKGW 132

RESULT 12
ADN99821
ID      ADN99821 standard; protein; 524 AA.
XX
XX      AC ADN99821;
XX
XX      DT 29-JUL-2004 (first entry)
XX
XX      DE Novel human protein sequence #637.
XX
XX      anti-inflammatory; dermatological; neuroprotective; immunomodulator;
KW      antibacterial; virucide; antiparasitic; cytostatic; gene therapy;
KW      vaccine; inflammatory; CNS; immune disorder; cancer; psoriasis; diabetes;
KW      early aging; hormonal imbalance; ischemic heart disease;
KW      ulcerative colitis.
XX
XX      OS Homo sapiens.

```


CC phenyl. The method is used for Radiolabeling peptide-containing targeting
 CC vectors such as proteins, antibodies, antibody fragments and receptor-
 CC targeted peptides for use in routine clinical positron emission
 CC tomography. The method is simple and efficient. The method uses the
 CC unique property of the free thiol groups which are rapidly alkylated at
 CC neutral pH and moderate temperature. Sequences AAY03714-716 represent
 CC examples of F-18 labeled peptides used in the method of detecting a
 CC tissue
 XX
 SQ Sequence 8 AA;
 Query Match 75.0%; Score 45; DB 2; Length 8;
 Best Local Similarity 85.7%; Pred. No. 1.8e+06;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3 YWGGGYW 9
 ||| |||
 Db 2 YWGGGYW 8
 RESULT 15
 AAY76817
 ID AAY76817 standard; peptide; 8 AA.
 XX
 AC AAY76817;
 XX
 DT 28-APR-2000 (first entry)
 XX
 DE Immunogenic peptide for bi-specific antibody recognition.
 XX
 KW Immunogenic peptide; bi-specific antibody; diagnosis; immune response;
 KW diseased tissue identification; therapy.
 XX
 OS Synthetic.
 XX
 FH Key Location/Qualifiers
 FT Misc-difference 1 /note= "acetylated; modified with free amino acid group,
 FT protected amino acid group, chelating agent or a metal-
 FT chelate complex"
 FT Misc-difference 2 /note= "D-form residue"
 FT Misc-difference 3 /note= "D-form residue"
 FT Misc-difference 5 /note= "D-form residue"
 FT /note= "D-form residue; modified with free amino acid
 FT group, protected amino acid group, chelating agent or a
 FT metal-chelate complex"
 FT Misc-difference 7 /note= "D-form residue"
 FT Misc-difference 8 /note= "D-form residue"
 FT /note= "D-form residue"
 XX
 DN W09966951-A2.
 XX
 XX 29-DEC-1999.
 XX
 XX 22-JUN-1999; 99WO-US013879.
 XX
 XX 22-JUN-1998; 98US-0090142P.
 PR 14-OCT-1998; 98US-0104156P.
 PR
 XX (IMMU-) IMMUNOMEDICS INC.
 XX
 XX Hansen HJ, Griffiths GL, Leung S, McBride WJ, Qu Z;
 XX WPI; 2000-160561/14.
 XX
 XX Bi-specific antibodies that bind specific target tissue and targeted
 PT conjugates.
 XX
 PS Claim 22; Page 61; 76pp; English.
 XX

CC This sequence represents an immunogenic peptide for a bi-specific
 CC antibody. The invention relates to a method of treating or identifying
 CC diseased tissues in a patient comprising administering a bi-specific
 CC antibody (or fragment) having at least 1 arm (A) that specifically
 CC binds a targeted tissue and at least 1 arm (B) that specifically binds a
 CC targetable conjugate. The methods and bi-specific antibodies and fusion
 CC proteins are useful for pre-targeting methods of diagnosis and therapy.
 CC It is advantageous to raise bi-specific antibodies against a targetable
 CC conjugate that is capable of carrying at least 1 diagnostic or
 CC therapeutic agent. The characteristics of the chelator, metal chelate
 CC complex, therapeutic agent or diagnostic agent can be varied to
 CC accommodate differing applications without raising new bi-specific
 CC antibodies for each new application. The targetable conjugate is selected
 CC to elicit sufficient immune responses and also for rapid in vivo
 CC clearance when used within the bi-specific antibody targeting method
 XX
 SQ Sequence 8 AA;
 Query Match 75.0%; Score 45; DB 3; Length 8;
 Best Local Similarity 85.7%; Pred. No. 1.8e+06;
 Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
 QY 3 YWGGGYW 9
 ||| |||
 Db 2 YWGGGYW 8
 Search completed: April 19, 2005, 06:39:26
 Job time : 179 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2005, 06:42:37 ; Search time 133 Seconds
(without alignments)
22.491 Million cell updates/sec

Title: SEQ2

Perfect score: 60

Sequence: 1 xgywgkgyw 9

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370683 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

- Published Applications AA:*
- 1: /cgn2_6/ptodata/2/pubpaa/US07_PUBCOMB.pep.*
 - 2: /cgn2_6/ptodata/2/pubpaa/PCT_NEW_PUB.pep.*
 - 3: /cgn2_6/ptodata/2/pubpaa/US06_NEW_PUB.pep.*
 - 4: /cgn2_6/ptodata/2/pubpaa/US06_PUBCOMB.pep.*
 - 5: /cgn2_6/ptodata/2/pubpaa/US07_NEW_PUB.pep.*
 - 6: /cgn2_6/ptodata/2/pubpaa/PCTUS_PUBCOMB.pep.*
 - 7: /cgn2_6/ptodata/2/pubpaa/US08_NEW_PUB.pep.*
 - 8: /cgn2_6/ptodata/2/pubpaa/US08_PUBCOMB.pep.*
 - 9: /cgn2_6/ptodata/2/pubpaa/US09A_PUBCOMB.pep.*
 - 10: /cgn2_6/ptodata/2/pubpaa/US09B_PUBCOMB.pep.*
 - 11: /cgn2_6/ptodata/2/pubpaa/US09C_PUBCOMB.pep.*
 - 12: /cgn2_6/ptodata/2/pubpaa/US09D_PUBCOMB.pep.*
 - 13: /cgn2_6/ptodata/2/pubpaa/US10A_PUBCOMB.pep.*
 - 14: /cgn2_6/ptodata/2/pubpaa/US10B_PUBCOMB.pep.*
 - 15: /cgn2_6/ptodata/2/pubpaa/US10C_PUBCOMB.pep.*
 - 16: /cgn2_6/ptodata/2/pubpaa/US10D_PUBCOMB.pep.*
 - 17: /cgn2_6/ptodata/2/pubpaa/US10E_PUBCOMB.pep.*
 - 18: /cgn2_6/ptodata/2/pubpaa/US11_NEW_PUB.pep.*
 - 19: /cgn2_6/ptodata/2/pubpaa/US60_NEW_PUB.pep.*
 - 20: /cgn2_6/ptodata/2/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	76.7	164	9 US-09-764-864-1230	Sequence 1230, Ap
2	44.5	74.2	440	15 US-10-282-122A-48322	Sequence 48322, A
3	43	71.7	446	15 US-10-282-122A-51216	Sequence 51216, A
4	43	71.7	1329	10 US-09-769-744A-186	Sequence 186, App
5	43	71.7	1329	17 US-10-472-928-4886	Sequence 4886, Ap
6	42	70.0	98	16 US-10-437-963-115321	Sequence 115321, A
7	42	70.0	141	16 US-10-767-701-36469	Sequence 36469, A
8	42	70.0	911	13 US-10-141-132-2	Sequence 2, Appli
9	41	68.3	50	16 US-10-742-379-136	Sequence 136, App
10	41	68.3	172	16 US-10-437-963-128022	Sequence 128022, A
11	41	68.3	445	15 US-10-282-122A-48093	Sequence 48093, A
12	41	68.3	453	15 US-10-425-114-64598	Sequence 64598, A
13	41	68.3	469	16 US-10-437-963-131752	Sequence 131752, A

14	68.3	1268	14	US-10-171-319-11	Sequence 11, Appl
15	67.5	247	15	US-10-424-599-184877	Sequence 184877, A
16	66.7	124	15	US-10-424-599-241189	Sequence 241189, A
17	66.7	171	15	US-10-425-114-48479	Sequence 48479, A
18	66.7	341	15	US-10-369-493-11021	Sequence 11021, A
19	66.7	608	13	US-10-029-009-11	Sequence 11, Appl
20	66.7	626	13	US-10-029-009-23	Sequence 23, Appl
21	66.7	729	14	US-10-156-761-12628	Sequence 12628, A
22	66.7	961	9	US-09-801-368-132	Sequence 132, App
23	65.0	48	16	US-10-437-963-174413	Sequence 174413, A
24	65.0	50	11	US-09-864-408A-5062	Sequence 5062, Ap
25	65.0	111	15	US-10-108-260A-2943	Sequence 2943, Ap
26	65.0	111	15	US-10-108-260A-3105	Sequence 3105, Ap
27	65.0	113	15	US-10-369-493-19264	Sequence 19264, A
28	65.0	115	15	US-10-424-599-151347	Sequence 151347, A
29	65.0	118	15	US-10-424-599-181122	Sequence 181122, A
30	65.0	120	17	US-10-891-972-23	Sequence 23, Appl
31	65.0	122	17	US-10-891-972-33	Sequence 33, Appl
32	65.0	138	9	US-09-796-744-15	Sequence 15, Appl
33	65.0	138	14	US-10-231-452-62	Sequence 62, Appl
34	65.0	186	17	US-10-472-928-1008	Sequence 1008, Ap
35	65.0	192	16	US-10-437-963-166130	Sequence 166130, A
36	65.0	249	10	US-09-880-748-1838	Sequence 1838, Ap
37	65.0	249	15	US-10-293-418-1838	Sequence 1838, Ap
38	65.0	336	14	US-10-156-761-13391	Sequence 13391, A
39	65.0	343	15	US-10-296-115-1186	Sequence 1186, Ap
40	65.0	394	16	US-10-437-963-146963	Sequence 146963, A
41	65.0	481	15	US-10-282-122A-51389	Sequence 51389, A
42	65.0	584	15	US-10-108-260A-3181	Sequence 3181, Ap
43	65.0	697	14	US-10-218-957-4	Sequence 4, Appli
44	65.0	697	15	US-10-072-012-511	Sequence 511, App
45	65.0	809	15	US-10-072-012-509	Sequence 509, App

ALIGNMENTS

RESULT 1

US-09-764-864-1230
; Sequence 1230, Application US/09764864
; Patent No. US20020132753A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT23
; CURRENT APPLICATION NUMBER: US/09/764,864
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - consult PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1792
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1230
; LENGTH: 164
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (2)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (32)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (34)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (84)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (119)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (138)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids

; NAME/KEY: SITE
; LOCATION: (140)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (143)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
; NAME/KEY: SITE
; LOCATION: (161)
; OTHER INFORMATION: Xaa equals any of the naturally occurring L-amino acids
US-09-764-864-1230

Query Match 76.7%; Score 46; DB 9; Length 164;
Best Local Similarity 66.7%; Pred. No. 38;
Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWKGKGYW 9
Db 124 SGHWKGEW 132

RESULT 2
US-10-282-122A-48322
; Sequence 48322, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48322
; LENGTH: 440
; TYPE: PRT
; ORGANISM: Bacteroides fragilis

US-10-282-122A-48322

Query Match 74.2%; Score 44.5; DB 15; Length 440;
Best Local Similarity 77.8%; Pred. No. 1.5e+02;
Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 XGYWKGKGYW 9
Db 211 PGWKG-GYW 218

RESULT 3
US-10-282-122A-51216
; Sequence 51216, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA 034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 51216
; LENGTH: 446
; TYPE: PRT
; ORGANISM: Bordetella pertussis

US-10-282-122A-51216

Query Match 71.7%; Score 43; DB 15; Length 446;
Best Local Similarity 55.6%; Pred. No. 2.5e+02;
Matches 5; Conservative 3; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWKGKGYW 9
Db 25 PGWKGGEW 33

RESULT 4
US-09-769-744A-166
; Sequence 166, Application US/09769744A
; Publication No. US20030134407A1
; GENERAL INFORMATION:
; APPLICANT: Le Page, Richard WF
; APPLICANT: Wells, Jeremy M
; APPLICANT: Hanniffy, Sean B

US-09-769-744A-166

; APPLICANT: Hansbro, Philip M
; TITLE OF INVENTION: Proteins
; FILE REFERENCE: PWC/P21122WO
; CURRENT APPLICATION NUMBER: US/09/769,744A
; PRIOR FILING DATE: 2001-01-26
; PRIOR APPLICATION NUMBER: PCT/GB99/02452
; PRIOR FILING DATE: 1999-07-27
; PRIOR APPLICATION NUMBER: GB 9816336.3
; PRIOR FILING DATE: 1998-07-27
; PRIOR APPLICATION NUMBER: US 60/125329
; PRIOR FILING DATE: 1999-03-19
; NUMBER OF SEQ ID NOS: 196
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 166
; LENGTH: 1329
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-769-744A-166

Query Match 71.7%; Score 43; DB 10; Length 1329;
Best Local Similarity 55.6%; Pred. No. 6.6e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 XGYWKGWY 9
Db 886 ISYWSKGW 894

RESULT 5
US-10-472-928-4826
; Sequence 4826, Application US/10472928
; Publication No. US20050020813A1
; GENERAL INFORMATION:
; APPLICANT: CHIRON SpA
; APPLICANT: THE INSTITUTE FOR GENOMIC RESEARCH
; TITLE OF INVENTION: STREPTOCOCCUS PNEUMONIAE PROTEINS AND NUCLEIC ACIDS
; FILE REFERENCE: P026926WO
; CURRENT APPLICATION NUMBER: US/10/472,928
; CURRENT FILING DATE: 2003-09-26
; PRIOR APPLICATION NUMBER: GB-0107658.7
; PRIOR FILING DATE: 2001-03-27
; NUMBER OF SEQ ID NOS: 4979
; SOFTWARE: SeqWin99, version 1.03
; SEQ ID NO 4826
; LENGTH: 1329
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; OTHER INFORMATION: beta-N-acetylhexosaminidase (strH)
; OTHER INFORMATION: Cellular location: Peptidoglycan-bound (LPXTG)
; OTHER INFORMATION: Feature of note: WY motif
; OTHER INFORMATION: Similar to strain R6 sequence 15902101 (0.E+01)
US-10-472-928-4826

Query Match 71.7%; Score 43; DB 17; Length 1329;
Best Local Similarity 55.6%; Pred. No. 6.6e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 XGYWKGWY 9
Db 886 ISYWSKGW 894

RESULT 6
US-10-437-963-115321
; Sequence 115321, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei

; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 115321
; LENGTH: 98
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_1892C.1.pap
US-10-437-963-115321

Query Match 70.0%; Score 42; DB 16; Length 98;
Best Local Similarity 66.7%; Pred. No. 87;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 XGYWKGWY 9
Db 1 WGRWKGAW 9

RESULT 7
US-10-767-701-36469
; Sequence 36469, Application US/10767701
; Publication No. US20040172684A1
; GENERAL INFORMATION:
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53535)B
; CURRENT APPLICATION NUMBER: US/10/767,701
; CURRENT FILING DATE: 2004-01-29
; NUMBER OF SEQ ID NOS: 63128
; SEQ ID NO 36469
; LENGTH: 141
; TYPE: PRT
; ORGANISM: Sorghum bicolor
; FEATURE:
; OTHER INFORMATION: Clone ID: SORBI-28MAY03-C9290_1.pap
US-10-767-701-36469

Query Match 70.0%; Score 42; DB 16; Length 141;
Best Local Similarity 66.7%; Pred. No. 1.2e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 XGYWKGWY 9
Db 36 HGRWKGAW 44

RESULT 8
US-10-141-132-2
; Sequence 2, Application US/10141132
; Publication No. US20020132331A1
; GENERAL INFORMATION:
; APPLICANT: Wei et al.
; TITLE OF INVENTION: Human DNA Ligase IV
; FILE REFERENCE: PF142PDI1
; CURRENT APPLICATION NUMBER: US/10/141,132
; CURRENT FILING DATE: 2002-05-09
; PRIOR APPLICATION NUMBER: 08/461,562
; PRIOR FILING DATE: 1995-06-05
; PRIOR APPLICATION NUMBER: PCT/US94/12922
; PRIOR FILING DATE: 1994-11-08
; NUMBER OF SEQ ID NOS: 9
; SOFTWARE: PatentIn version 3.1

```
; SEQ ID NO 2
; LENGTH: 911
; TYPE: PRT
; ORGANISM: human
US-10-141-132-2

Query Match          70.0%; Score 42; DB 13; Length 911;
Best Local Similarity 85.7%; Pred. No. 6.5e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY 1 XGYWGKG 7
   :|||||
Db 468 GGYWGKG 474

RESULT 9
US-10-742-379-136
; Sequence 136, Application US/10742379
; Publication No. US20040181033A1
; GENERAL INFORMATION:
; APPLICANT: Han, HQ
; APPLICANT: Min, Hosung
; APPLICANT: Boone, Thomas Charles
; TITLE OF INVENTION: BINDING AGENTS WHICH INHIBIT MYOSTATIN
; FILE REFERENCE: A-828 (US)
; CURRENT APPLICATION NUMBER: US/10/742,379
; CURRENT FILING DATE: 2003-12-19
; PRIOR APPLICATION NUMBER: US 60/435,923
; PRIOR FILING DATE: 2002-12-20
; NUMBER OF SEQ ID NOS: 634
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 136
; LENGTH: 50
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Myostatin Binding Peptide
US-10-742-379-136

Query Match          58.3%; Score 41; DB 16; Length 50;
Best Local Similarity 55.6%; Pred. No. 66;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWGKG 9
   :|||
Db 35 TGHWACGYW 43

RESULT 10
US-10-437-963-128022
; Sequence 128022, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 128022
; LENGTH: 172
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure

; LOCATION: (1)..(172)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_30416C.1.pap
US-10-437-963-128022

Query Match          68.3%; Score 41; DB 16; Length 172;
Best Local Similarity 55.6%; Pred. No. 2e+02;
Matches 5; Conservative 2; Mismatches 0; Indels 0; Gaps 0;

QY 1 XGYWGKG 9
   :|||
Db 36 GGMWGRQYW 44

RESULT 11
US-10-282-122A-48093
; Sequence 48093, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 48093
; LENGTH: 445
; TYPE: PRT
; ORGANISM: Burkholderia cepacia
US-10-282-122A-48093

Query Match          68.3%; Score 41; DB 15; Length 445;
Best Local Similarity 75.0%; Pred. No. 4.7e+02;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWGKG 8
   :|||
Db 137 AGYWKKG 144
```

RESULT 12
US-10-425-114-64598
; Sequence 64598, Application US/10425114
; Publication No. US2004003488A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 64598
; LENGTH: 453
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB4172-021-C8_FLI.pep
US-10-425-114-64598

Query Match 68.3%; Score 41; DB 15; Length 453;
Best Local Similarity 55.6%; Pred. No. 4.8e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 XGYWKGKGYW 9
: |||||
Db 33 HGMWGRHYW 41

RESULT 13
US-10-437-963-131752
; Sequence 131752, Application US/10437963
; Publication No. US2004012343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 131752
; LENGTH: 469
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_33789C.1.pep
US-10-437-963-131752

Query Match 68.3%; Score 41; DB 16; Length 469;
Best Local Similarity 66.7%; Pred. No. 5e+02;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

Qy 1 XGYWKGKGYW 9
: |||||
Db 199 GSYGKGKGYW 207

RESULT 14
US-10-171-319-11

; Sequence 11, Application US/10171319
; Publication No. US20030157633A1
; GENERAL INFORMATION:
; APPLICANT: Ardem Patapoutian
; APPLICANT: Andrea Peier
; APPLICANT: Peter McIntyre
; APPLICANT: Stuart Bevan
; APPLICANT: Chuanzheng Song
; APPLICANT: Pamposh Ganju
; TITLE OF INVENTION: VANILLOID RECEPTOR-RELATED NUCLEIC ACIDS
; TITLE OF INVENTION: AND POLYPEPTIDES
; FILE REFERENCE: 4-32048A
; CURRENT APPLICATION NUMBER: US/10/171,319
; CURRENT FILING DATE: 2002-10-24
; PRIOR APPLICATION NUMBER: 60/297,835
; PRIOR FILING DATE: 2001-06-13
; PRIOR APPLICATION NUMBER: 60/351,238
; PRIOR FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: 60/352,914
; PRIOR FILING DATE: 2002-01-29
; PRIOR APPLICATION NUMBER: 60/357,161
; PRIOR FILING DATE: 2002-02-12
; PRIOR APPLICATION NUMBER: 60/381,086
; PRIOR FILING DATE: 2002-05-15
; PRIOR APPLICATION NUMBER: 60/381,739
; PRIOR FILING DATE: 2002-05-16
; NUMBER OF SEQ ID NOS: 114
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 11
; LENGTH: 1268
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-171-319-11

Query Match 68.3%; Score 41; DB 14; Length 1268;
Best Local Similarity 55.8%; Pred. No. 1.2e+03;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

Qy 1 XGYWKGKGYW 9
: |||||
Db 152 GGVWGRGLW 160

RESULT 15
US-10-424-599-184877
; Sequence 184877, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 184877
; LENGTH: 247
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(247)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_13795C.1.pep
US-10-424-599-184877

Query Match 67.5%; Score 40.5; DB 15; Length 247;
Best Local Similarity 66.7%; Pred. No. 3.3e+02;
Matches 6; Conservative 2; Mismatches 0; Indels 1; Gaps 1;

Qy 1 XGYWKGYY 9
:||||: ||
Db 108 DGYGR-YW 115

Search completed: April 19, 2005, 06:55:18
Job time : 133 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: April 19, 2005, 06:36:32 ; Search time 41 Seconds
(without alignments)
16.386 Million cell updates/sec

Title: SEQ2

Perfect score: 60

Sequence: 1 xgywkggyw 9

Scoring table: BLOSUM62DX

Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*

1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*

2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*

3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*

4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*

5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*

6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	80.0	1365	3	US-09-376-330-18
2	44.5	74.2	339	4	US-09-252-991A-26841
3	44	73.3	383	4	US-09-252-991A-29457
4	43	71.7	1290	4	US-09-107-433-4399
5	43	71.7	1303	4	US-09-583-110-5037
6	42	70.0	844	4	US-09-341-505-2
7	42	70.0	911	4	US-08-461-562B-2
8	42	70.0	922	4	US-09-902-540-12187
9	41	68.3	142	2	US-08-860-174A-7
10	41	68.3	274	2	US-08-860-174A-12
11	41	68.3	282	2	US-08-860-174A-10
12	41	68.3	320	4	US-09-248-796A-14837
13	41	68.3	1447	3	US-09-376-330-17
14	40	66.7	139	4	US-09-471-276-837
15	40	66.7	588	4	US-09-540-236-2481
16	40	66.7	624	3	US-08-947-965-78
17	40	66.7	655	1	US-08-469-202-27
18	40	66.7	655	2	US-08-469-202-28
19	40	66.7	655	1	US-08-484-434C-34
20	40	66.7	655	2	US-08-484-434C-35
21	40	66.7	655	4	US-09-384-361-34
22	40	66.7	655	4	US-09-384-361-35
23	39	65.0	52	4	US-09-513-999C-4456
24	39	65.0	100	4	US-09-513-999C-5031
25	39	65.0	126	4	US-09-902-540-15783
26	39	65.0	185	4	US-09-583-110-2898
27	39	65.0	196	4	US-09-107-433-5167

ALIGNMENTS

RESULT 1

US-09-376-330-18

; Sequence 18, Application US/09376330

; Patent No. 6399321

; GENERAL INFORMATION:

; APPLICANT: Tessier, Daniel C.

; APPLICANT: Dignard, Daniel

; APPLICANT: Bergeron, John J.M.

; APPLICANT: Thomas, David Y.

; TITLE OF INVENTION: Method for screening for

; TITLE OF INVENTION: UDP-glucose:glycoprotein glucosyltransferase (UGGT) activity

; TITLE OF INVENTION: and nucleic acid encoding for UGGT

; FILE REFERENCE: 2139-9"US"

; CURRENT APPLICATION NUMBER: US/09/376,330

; CURRENT FILING DATE: 1999-08-18

; NUMBER OF SEQ ID NOS: 35

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 18

; LENGTH: 1365

; TYPE: PRT

; ORGANISM: Artificial Sequence

; FEATURE:

; OTHER INFORMATION: S. cerevisiae KRE5

US-09-376-330-18

Query Match 80.0%; Score 48; DB 3; Length 1365;

Best Local Similarity 66.7%; Pred. NO. 48;

Matches 6; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 XGYWKGYYW 9

Db 1233 DGYWKGYYW 1241

RESULT 2

US-09-252-991A-26841

; Sequence 26841, Application US/09252991A

; Patent No. 6551795

; GENERAL INFORMATION:

; APPLICANT: Marc J. Rubenfield et al.

; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS

; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS

; FILE REFERENCE: 107196.136

; CURRENT APPLICATION NUMBER: US/09/252,991A

; CURRENT FILING DATE: 1999-02-18

; PRIOR APPLICATION NUMBER: US 60/074,788

; PRIOR FILING DATE: 1998-02-18

; PRIOR APPLICATION NUMBER: US 60/094,190

; PRIOR FILING DATE: 1998-07-27

; NUMBER OF SEQ ID NOS: 33142

Sequence 5005, Ap
Sequence 21708, A
Patent No. 5194600
Patent No. 5194600
Sequence 4506, Ap
Sequence 9, Appli
Sequence 9, Appli
Sequence 6122, Ap
Sequence 11644, A
Sequence 12, Appl
Sequence 99, Appl
Sequence 99, Appl
Sequence 296, App
Sequence 144, App
Sequence 99, Appl
Sequence 99, Appl
Sequence 99, Appl
Sequence 99, Appl

```
; SEQ ID NO 26841
; LENGTH: 339
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-26841

Query Match          74.2%; Score 44.5; DB 4; Length 339;
Best Local Similarity 77.8%; Pred. No. 41;
Matches 7; Conservative 1; Mismatches 0; Indels 1; Gaps 1;

QY 1 XGYWKGWY 9
   :||| |||
Db 260 GGYWG-GYW 267

RESULT 3
US-09-252-991A-29457
; Sequence 29457, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252.991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 29457
; LENGTH: 383
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-29457

Query Match          73.3%; Score 44; DB 4; Length 383;
Best Local Similarity 66.7%; Pred. No. 54;
Matches 6; Conservative 1; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWKGWY 9
   :||| |||
Db 222 VGTWGGGYW 230

RESULT 4
US-09-107-433-4399
; Sequence 4399, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGN
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
```

```
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4399:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1290 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...1290
; SEQUENCE DESCRIPTION: SEQ ID NO: 4399:
US-09-107-433-4399

Query Match          71.7%; Score 43; DB 4; Length 1290;
Best Local Similarity 55.6%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWKGWY 9
   :||| |||
Db 847 ISYWSKGWY 855

RESULT 5
US-09-583-110-5037
; Sequence 5037, Application US/09583110
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; TITLE OF INVENTION: Pneumoniae for Diagnostics and Therapeutics
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; CURRENT FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 5037
; LENGTH: 1303
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-5037

Query Match          71.7%; Score 43; DB 4; Length 1303;
Best Local Similarity 55.6%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWKGWY 9
   :||| |||
Db 860 ISYWSKGWY 868

RESULT 6
US-09-341-505-2
; Sequence 2, Application US/09341505
; Patent No. 6753158
; GENERAL INFORMATION:
; APPLICANT: Jackson, Stephen P
```



```
; APPLICANT: Critchlow, Susan E
; TITLE OF INVENTION: Assays, agents, therapy and diagnosis relating to
; TITLE OF INVENTION: modulation of cellular DNA repair activity
; FILE REFERENCE: MEWE-005
; CURRENT APPLICATION NUMBER: US/09/341,505
; CURRENT FILING DATE: 1999-07-12
; EARLIER APPLICATION NUMBER: PCT/GB98/00095
; EARLIER FILING DATE: 1998-01-13
; EARLIER APPLICATION NUMBER: GB 9700574.8
; EARLIER FILING DATE: 1997-01-13
; EARLIER APPLICATION NUMBER: GB 9713131.2
; EARLIER FILING DATE: 1997-06-20
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in Ver. 2.1
; SEQ ID NO 2
; LENGTH: 844
; TYPE: PRT
; ORGANISM: Homo sapiens
; US-09-341-505-2

Query Match      70.0%; Score 42; DB 4; Length 844;
Best Local Similarity 85.7%; Pred. No. 2.2e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 XGYWKGK 7
Db      401 GGYWKGK 407

RESULT 7
US-08-461-562B-2
; Sequence 2, Application US/08461562B
; Patent No. 6455274
; GENERAL INFORMATION:
; APPLICANT: WEI, YING-FEI
; APPLICANT: HASELTINE, WILLIAM H
; TITLE OF INVENTION: HUMAN DNA LIGASE IV
; NUMBER OF SEQUENCES: 9
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: HUMAN GENOME SCIENCES, INC.
; STREET: 9410 KEY WEST AVE
; CITY: ROCKVILLE
; STATE: MD
; COUNTRY: US
; ZIP: 20850
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/461,562B
; FILING DATE: 05-JUN-1995
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US94/12922
; FILING DATE: 08-NOV-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: A. ANDERS BROOKES
; REGISTRATION NUMBER: 36,373
; REFERENCE/DOCKET NUMBER: PF142P1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 301-309-8504
; TELEFAX: 301-309-8439
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 911 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-461-562B-2

Query Match      70.0%; Score 42; DB 4; Length 911;
```

```
Best Local Similarity 85.7%; Pred. No. 2.3e+02;
Matches 6; Conservative 1; Mismatches 0; Indels 0; Gaps 0;

QY      1 XGYWKGK 7
Db      468 GGYWKGK 474

RESULT 8
US-09-902-540-12187
; Sequence 12187, Application US/09902540
; Patent No. 6833447
; GENERAL INFORMATION:
; APPLICANT: Goldman, Barry S.
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Wiegand, Roger C.
; TITLE OF INVENTION: Myxococcus xanthus Genome Sequences and Uses Thereof
; FILE REFERENCE: 38-10(15849)B
; CURRENT APPLICATION NUMBER: US/09/902,540
; CURRENT FILING DATE: 2001-07-10
; PRIOR APPLICATION NUMBER: 60/217,883
; PRIOR FILING DATE: 2000-07-10
; NUMBER OF SEQ ID NOS: 16825
; SEQ ID NO 12187
; LENGTH: 922
; TYPE: PRT
; ORGANISM: Myxococcus xanthus
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(922)
; OTHER INFORMATION: unsure at all Xaa locations
US-09-902-540-12187

Query Match      70.0%; Score 42; DB 4; Length 922;
Best Local Similarity 55.6%; Pred. No. 2.4e+02;
Matches 5; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 XGYWKGK 9
Db      730 GGYWGENYW 738

RESULT 9
US-08-860-174A-7
; Sequence 7, Application US/08860174A
; Patent No. 5989830
; GENERAL INFORMATION:
; APPLICANT: DAVIS, Paul James
; APPLICANT: VAN DER LOGT, Cornelis Paul Erik
; APPLICANT: VERHOEIJEN, Martine Elisa
; APPLICANT: WILSON, Steve
; TITLE OF INVENTION: A BIFUNCTIONAL OR BIVALENT
; TITLE OF INVENTION: ANTIBODY FRAGMENT ANALOGUE
; NUMBER OF SEQUENCES: 31
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: PILLSBURY, MADISON & SUTRO LLP
; STREET: 1100 New York Avenue, N.W.
; STREET: 9th Floor, East Tower
; CITY: WASHINGTON, D.C.
; STATE:
; COUNTRY: UNITED STATES
; ZIP: 20005-3918
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: MS WORD, WINDOWS 95, Ver. 4.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/860,174A
; FILING DATE: June 16, 1997
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: EP 95307332.7
```

/ FILING DATE: October 16, 1995
/ APPLICATION NUMBER: PCT/EP/96/03605
/ FILING DATE: August 14, 1996
/ INFORMATION FOR SEQ ID NO: 7:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 142 amino acids
/ TYPE: amino acid
/ TOPOLOGY: linear
/ MOLECULE TYPE: protein
/ US-08-860-174A-7

Query Match 68.3%; Score 41; DB 2; Length 142;
Best Local Similarity 66.7%; Pred. No. 57;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 XGYWGKGYW 9
Db 121 HGYGKGYF 129

RESULT 10

US-08-860-174A-12
/ Sequence 12, Application US/08860174A

/ Patent No. 5989830

/ GENERAL INFORMATION:

/ APPLICANT: DAVIS, Paul James
/ APPLICANT: VAN DER LOGT, Cornelis Paul Erik
/ APPLICANT: VERHOEIJEN, Martine Elisa

/ APPLICANT: WILSON, Steve

/ TITLE OF INVENTION: A BIFUNCTIONAL OR BIVALENT

/ TITLE OF INVENTION: ANTIBODY FRAGMENT ANALOGUE

/ NUMBER OF SEQUENCES: 31

/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: PILLSBURY, MADISON & SUTRO LLP

/ STREET: 1100 New York Avenue, N.W.

/ STREET: 9th Floor, East Tower

/ CITY: WASHINGTON, D.C.

/ STATE:

/ COUNTRY: UNITED STATES

/ ZIP: 20005-3918

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS DOS Text

/ SOFTWARE: MS WORD, WINDOWS 95, Ver. 4.0

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/860,174A

/ FILING DATE: June 16, 1997

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: EP 95307332.7

/ FILING DATE: October 16, 1995

/ APPLICATION NUMBER: PCT/EP/96/03605

/ FILING DATE: August 14, 1996

/ INFORMATION FOR SEQ ID NO: 12:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 274 amino acids

/ TYPE: amino acid

/ TOPOLOGY: linear

/ MOLECULE TYPE: protein

/ US-08-860-174A-12

Query Match 68.3%; Score 41; DB 2; Length 274;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 XGYWGKGYW 9
Db 121 HGYGKGYF 129

RESULT 11

US-08-860-174A-10

/ Sequence 10, Application US/08860174A

/ Patent No. 5989830
/ GENERAL INFORMATION:
/ APPLICANT: DAVIS, Paul James
/ APPLICANT: VAN DER LOGT, Cornelis Paul Erik
/ APPLICANT: VERHOEIJEN, Martine Elisa
/ APPLICANT: WILSON, Steve
/ TITLE OF INVENTION: A BIFUNCTIONAL OR BIVALENT
/ TITLE OF INVENTION: ANTIBODY FRAGMENT ANALOGUE
/ NUMBER OF SEQUENCES: 31
/ CORRESPONDENCE ADDRESS:

/ ADDRESSEE: PILLSBURY, MADISON & SUTRO LLP

/ STREET: 1100 New York Avenue, N.W.

/ STREET: 9th Floor, East Tower

/ CITY: WASHINGTON, D.C.

/ STATE:

/ COUNTRY: UNITED STATES

/ ZIP: 20005-3918

/ COMPUTER READABLE FORM:

/ MEDIUM TYPE: Floppy disk

/ COMPUTER: IBM PC compatible

/ OPERATING SYSTEM: PC-DOS/MS-DOS DOS Text

/ SOFTWARE: MS WORD, WINDOWS 95, Ver. 4.0

/ CURRENT APPLICATION DATA:

/ APPLICATION NUMBER: US/08/860,174A

/ FILING DATE: June 16, 1997

/ PRIOR APPLICATION DATA:

/ APPLICATION NUMBER: EP 95307332.7

/ FILING DATE: October 16, 1995

/ APPLICATION NUMBER: PCT/EP/96/03605

/ FILING DATE: August 14, 1996

/ INFORMATION FOR SEQ ID NO: 10:

/ SEQUENCE CHARACTERISTICS:

/ LENGTH: 282 amino acids

/ TYPE: amino acid

/ TOPOLOGY: linear

/ MOLECULE TYPE: protein

/ US-08-860-174A-10

Query Match 68.3%; Score 41; DB 2; Length 282;
Best Local Similarity 66.7%; Pred. No. 1.1e+02;
Matches 6; Conservative 3; Mismatches 0; Indels 0; Gaps 0;

QY 1 XGYWGKGYW 9
Db 121 HGYGKGYF 129

RESULT 12

US-09-248-796A-14837

/ Sequence 14837, Application US/09248796A

/ Patent No. 6747137

/ GENERAL INFORMATION:

/ APPLICANT: Keith Weinstock et al

/ TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO CANDIDA ALBICANS

/ TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS

/ FILE REFERENCE: 107196.132

/ CURRENT APPLICATION NUMBER: US/09/248,796A

/ CURRENT FILING DATE: 1999-02-12

/ PRIOR APPLICATION NUMBER: US 60/074,725

/ PRIOR FILING DATE: 1998-02-13

/ PRIOR APPLICATION NUMBER: US 60/096,409

/ PRIOR FILING DATE: 1998-08-13

/ NUMBER OF SEQ ID NOS: 28208

/ SEQ ID NO 14837

/ LENGTH: 320

/ TYPE: PRT

/ ORGANISM: Candida albicans

/ US-09-248-796A-14837

Query Match 68.3%; Score 41; DB 4; Length 320;
Best Local Similarity 55.6%; Pred. No. 1.2e+02;
Matches 5; Conservative 4; Mismatches 0; Indels 0; Gaps 0;

QY 1 XGYWKGKGYW 9
Db 287 LGWFGKGF 295

RESULT 13
US-09-376-330-17
; Sequence 17, Application US/09376330
; Patent No. 6399321
; GENERAL INFORMATION:
; APPLICANT: Tessier, Daniel C.
; APPLICANT: Dignard, Daniel
; APPLICANT: Bergeron, John J.M.
; APPLICANT: Thomas, David Y.
; TITLE OF INVENTION: Method for screening for
; TITLE OF INVENTION: UDP-glucose:glycoprotein glucosyltransferase (UGGT) activity
; TITLE OF INVENTION: and nucleic acid encoding for UGGT
; FILE REFERENCE: 2139-9"US"
; CURRENT APPLICATION NUMBER: US/09/376,330
; CURRENT FILING DATE: 1999-08-18
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 17
; LENGTH: 1447
; TYPE: PRT
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: S. pombe UGGT
US-09-376-330-17

Query Match 68.3%; Score 41; DB 3; Length 1447;
Best Local Similarity 71.4%; Pred. No. 5e+02;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 YWKGKGYW 9
Db 1297 FWKKGW 1303

RESULT 14
US-09-471-276-837
; Sequence 837, Application US/09471276
; Patent No. 6822072
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Duclert A.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: Expressed Sequence Tags and Encoded Human Proteins.
; Patent No. 6822072
; FILE REFERENCE: GENSET.025CP1
; CURRENT APPLICATION NUMBER: US/09/471,276
; CURRENT FILING DATE: 1999-12-21
; EARLIER APPLICATION NUMBER: 09/057,719
; EARLIER FILING DATE: 1998-04-09
; EARLIER APPLICATION NUMBER: 09/069,047
; EARLIER FILING DATE: 1998-04-28
; EARLIER APPLICATION NUMBER: PCT/IB99/00712
; EARLIER FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 1622
; SOFTWARE: Patent.pm
; SEQ ID NO 837
; LENGTH: 139
; TYPE: PRT
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SIGNAL
; LOCATION: -19..-1
US-09-471-276-837

Query Match 66.7%; Score 40; DB 4; Length 139;
Best Local Similarity 71.4%; Pred. No. 77;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 3 YWKGKGYW 9
Db 132 YWQGTW 138

RESULT 15
US-09-540-236-2481
; Sequence 2481, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO MORAXELLA CATA
; TITLE OF INVENTION: FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2481
; LENGTH: 588
; TYPE: PRT
; ORGANISM: M.cattarrhalis
US-09-540-236-2481

Query Match 66.7%; Score 40; DB 4; Length 588;
Best Local Similarity 44.4%; Pred. No. 3e+02;
Matches 4; Conservative 3; Mismatches 2; Indels 0; Gaps 0;

QY 1 XGYWKGKGYW 9
Db 30 NSFWSKGF 38

Search completed: April 19, 2005, 06:44:44
Job time : 42 secs

This Page Blank (uspto)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 19:35:54 ; Search time 15.4419 Seconds
(without alignments)
49.847 Million cell updates/sec

Title: SEQ5
Perfect score: 66
Sequence: 1 cywgcgyw 8
Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 96216763 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : PIR 79:***
1: pir1:***
2: pir2:***
3: pir3:***
4: pir4:***

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	46	69.7	554	A28716	saposin precursor
2	45.5	68.9	269	E96704	hypothetical prote
3	45.5	68.9	283	H84430	hypothetical prote
4	42	63.6	48	A49739	relaxin - horse (f
5	42	63.6	143	I47053	relaxin B,C and A
6	42	63.6	342	T31757	hypothetical prote
7	42	63.6	347	T31755	hypothetical prote
8	42	63.6	351	T31758	hypothetical prote
9	42	63.6	473	I54210	N-acetylglactosam
10	42	63.6	533	KJHUAB	probable glycogen
11	42	63.6	863	D70770	hypothetical prote
12	41	62.1	83	B72392	late embryogenesis
13	41	62.1	280	G84839	muscarinic acetyl
14	41	62.1	441	H96368	muscarinic acetyl
15	41	62.1	589	A29476	muscarinic acetyl
16	41	62.1	599	B29514	muscarinic acetyl
17	41	62.1	590	S01114	muscarinic acetyl
18	41	62.1	590	S47572	muscarinic acetyl
19	41	62.1	590	AF1810	hypothetical prote
20	41	62.1	741	S73827	hypothetical prote
21	41	62.1	2970	T08839	polyprotein - marm
22	40	60.6	113	PH1663	Ig heavy chain V r
23	40	60.6	130	A69894	hypothetical prote
24	40	60.6	130	T44809	hypothetical prote
25	40	60.6	396	T04561	hypothetical prote
26	40	60.6	415	H86204	probable phosphos
27	40	60.6	441	F83694	short-chain fatty
28	40	60.6	456	T40367	n-acetylglucosam
29	40	60.6	499	T36462	hypothetical prote

30	40	60.6	527	1	SAHUP	saposin precursor
31	40	60.6	557	1	JH0604	saposin precursor
32	40	60.6	573	1	S33212	INDAI protein - fu
33	40	60.6	630	2	JC5374	angiotensin-conver
34	39.5	59.8	62	1	K3SHI6	major high-(glycin
35	39	59.1	116	2	H84004	hypothetical prote
36	39	59.1	118	1	A59846	conserved hypotet
37	39	59.1	163	2	H72759	probable dCTP deam
38	39	59.1	270	1	G83242	probable phosphos
39	39	59.1	304	2	S50136	nitrogenase (BC 1.
40	39	59.1	317	2	T38935	hypothetical prote
41	39	59.1	332	2	AF3304	oxidoreductase (BC
42	39	59.1	400	2	A39254	inositol-1,4-bisph
43	39	59.1	466	2	T05664	hypothetical prote
44	39	59.1	480	1	NIAIMA	nitrogenase (BC 1.
45	39	59.1	497	2	AF1986	nitrogenase molybd

ALIGNMENTS

RESULT 1

A28716
saposin precursor - rat
N:Alternate names: cerebroside sulfate activator protein; co-beta-glucosidase; component
ein (SAP); sphingolipid activator protein A2; sulfated glycoprotein 1; sulfate sulfat
N:Contains: prosaposin; saposin A; saposin B; saposin C; saposin D
C:Species: Rattus norvegicus (Norway rat)
C:Date: 10-Sep-1999 #sequence_revision 10-Sep-1999 #text_change 09-Jul-2004
C:Accession: A28716
R:Collard, M.W.; Sylvester, S.R.; Tsuruta, J.K.; Griswold, M.D.
Biochemistry 27, 4557-4564, 1988
A:Title: Biosynthesis and molecular cloning of sulfated glycoprotein 1 secreted by rat
A:Reference number: A28716; MUID:89000647; PMID:3048385
A:Accession: A28716
A:Molecule type: mRNA
A:Residues: 1-554 <COI>
A:Cross-references: UNIPROT:P10960; GB:M19936; NID:G206904; PIDN:AAA42136.1; PID:G20690
A:Note: parts of this sequence, including the amino end of the mature protein, were det
C:Function:
A:Description: saposins bind sphingolipids, form hydrophilic complexes and make them ac
A:Pathway: sphingolipid catabolism
A:Note: saposins A and C (SAP-2) activate hydrolysis of glucocerebroside by beta-glucos
A:Note: saposin B (SAP-1) activates hydrolysis of galactocerebroside sulfate by arylsul
A:Note: saposin D activates hydrolysis of sphingomyelin by sphingomyelin phosphodiester
C:Superfamily: saposin; saposin repeat homology
C:Keywords: alternative splicing; glycoprotein; lysosomal storage disease; lysosome; spi
F:1-16/Domain: signal sequence #status predicted <SIG>
F:17-554/Product: prosaposin #status predicted <PRO>
F:55-148/Domain: saposin repeat homology <SAP1>
F:60-143/Product: saposin A #status predicted <SAPA>
F:189-280/Domain: saposin repeat homology <SAP2>
F:194-273/Product: saposin B #status predicted <SAB1>
F:306-397/Domain: saposin repeat homology <SAP3>
F:310-389/Product: saposin C #status predicted <SAPC>
F:431-522/Domain: saposin repeat homology <SAP4>
F:437-514/Product: saposin D #status predicted <SAPD>
F:63-138,66-132,94-106,439-512,442-506,470-481/Disulfide bonds: #status predicted
F:80,214,331,456/Binding site: carboxydrate (Asn) (covalent) #status predicted
F:197-270,200-264,229-240,314-387,317-381,345-356/Disulfide bonds: #status predicted

Query Match 69.7%; Score 46; DB 1; Length 554;
Best Local Similarity 75.0%; Pred. No. 18;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CYWGCYW 8
| | | | |
Db 525 CWGPGYW 532

RESULT 2

E96704
hypothetical protein T23K23.3 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)
 C:Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C:Accession: E96704
 R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.
 Nature 408, 816-820, 2000
 A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.A.; Li, J.H.; Lin, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maity, R.; Marziani, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.
 A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
 A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
 A:Reference number: A86141; MUID:21016719; PMID:11130712
 A:Accession: E96704
 A:Status: Preliminary
 A:Molecule type: DNA
 A:Residues: 1-269 <STO>
 A:Cross-references: UNIPROT:Q9C9X6; GB:AE005173; NID:G6553884; PIDN:AAF16550.1; GSPDB:GN
 C:Genetics:
 A:Gene: T23K23.3
 A:Map position: 1

Query Match 68.9%; Score 45.5; DB 2; Length 269;

Best Local Similarity 77.8%; Pred. No. 12;
 Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CY-WGCCYW 8

DB 190 CYRWGCGW 198

RESULT 3

H84430

hypothetical protein At2g01930 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004

C:Accession: H84430

R:Lin, X.; Kaul, S.; Rounslev, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;

M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; VanAken, S.E.; Umayam, L.; Tallon, L.

eus, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J.

Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: H84430

A:Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-283 <STO>

A:Cross-references: UNIPROT:Q9SKD0; GB:AE002093; NID:G6598329; PIDN:AAF18588.1; GSPDB:GN

C:Genetics:

A:Gene: At2g01930

A:Map position: 2

Query Match

Best Local Similarity 68.9%; Score 45.5; DB 2; Length 283;

Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CY-WGCCYW 8

DB 204 CYRWGCGW 212

RESULT 4

A49739

relaxin - horse (fragments)

C:Species: Equus caballus (domestic horse)

C:Date: 07-Apr-1994 #sequence_revision 26-May-1995 #text_change 09-Jul-2004

C:Accession: B49739; A49739

R:Stewart, D.R.; Nevins, B.; Hadas, E.; Vandlen, R.

Endocrinology 129, 375-383, 1991

A:Title: Affinity purification and sequence determination of equine relaxin.

A:Reference number: A49739; MUID:91275796; PMID:2055195

A:Accession: B49739

A:Molecule type: protein
 A:Residues: 1-28 <STE>
 A:Cross-references: UNIPROT:P22969
 A:Accession: A49739
 A:Molecule type: protein
 A:Residues: 29-48 <ST2>
 C:Superfamily: insulin
 C:Keywords: hormone; pyroglutamic acid
 F:1-28/29-48/Product: relaxin #status experimental <MAT>
 F:1-28/Domain: chain B #status experimental <CHB>
 F:29-48/Domain: chain A #status experimental <CHA>
 F:1,29/Modified site: pyroglutamic acid (Gln) (in mature form) #status experimental
 F:10-34,22-39,35-48/Disulfide bonds: #status predicted

Query Match 63.6%; Score 42; DB 2; Length 48;

Best Local Similarity 100.0%; Pred. No. 10;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CYWGC 5

DB 35 CYWGC 39

RESULT 5

I47053

relaxin B,C and A chains - horse (fragment)

C:Species: Equus sp.

C:Date: 04-Sep-1997 #sequence_revision 07-Nov-1997 #text_change 16-Jul-1999

C:Accession: I47053

R:Klonisch, T.; Ryan, P.L.; Yamashiro, S.; Porter, D.G.

Biol. Reprod. 52, 1307-1315, 1995

A:Title: Partial complementary deoxyribonucleic acid cloning of equine relaxin messenger

A:Reference number: I47053; MUID:95359320; PMID:7543295

A:Accession: I47053

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: mRNA

A:Residues: 1-143 <KLO>

A:Cross-references: GB:S78800; NID:g1042059; PIDN:AAB35036.1; PID:g1042060

C:Superfamily: insulin

Query Match

Best Local Similarity 63.6%; Score 42; DB 2; Length 143;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CYWGC 5

DB 138 CYWGC 142

RESULT 6

T31757

hypothetical protein C07G3.4 - Caenorhabditis elegans

C:Species: Caenorhabditis elegans

C:Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004

C:Accession: T31757

R:Geisel, C.; Wamsley, P.

submitted to the EMBL Data Library, July 1997

A:Description: The sequence of C. elegans cosmid C07G3.

A:Reference number: Z21080

A:Accession: T31757

A:Status: Preliminary; translated from GB/EMBL/DBJ

A:Molecule type: DNA

A:Residues: 1-342 <GEI>

A:Cross-references: UNIPROT:O16325; EMBL:AF016432; PIDN:AAB65382.1; GSPDB:GN00023; CESP:

A:Experimental source: strain Bristol N2; clone C07G3

C:Genetics:

A:Gene: CESP:C07G3.4

A:Map position: 5

A:Introns: 66/3; 122/2; 167/2; 214/3; 319/1

Query Match

Best Local Similarity 63.6%; Score 42; DB 2; Length 342;

Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CYWGC 5
|||||
Db 97 CYWGC 101

RESULT 7

T31755
hypothetical protein C07G3.6 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31755
R:Geisler, C.; Wamsley, P.
submitted to the EMBL data Library, July 1997
A:Description: The sequence of *C. elegans* cosmid C07G3.
A:Reference number: Z21080
A:Accession: T31755
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-347 <GEI>
A:Cross-references: UNIPROT:O16323; EMBL:AF016432; PIDN:AAB65378.1; GSPDB:GN000023; CESP:
A:Experimental source: strain Bristol N2; clone C07G3
C:Genetics:
A:Gene: CESP:C07G3.6
A:Map position: 5
A:Introns: 66/3; 122/2; 167/2; 214/3; 319/1

Query Match 63.6%; Score 42; DB 2; Length 347;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CYWGC 5
|||||
Db 97 CYWGC 101

RESULT 8

T31758
hypothetical protein C07G3.3 - *Caenorhabditis elegans*
C:Species: *Caenorhabditis elegans*
C>Date: 29-Oct-1999 #sequence_revision 29-Oct-1999 #text_change 09-Jul-2004
C:Accession: T31758
R:Geisler, C.; Wamsley, P.
submitted to the EMBL data Library, July 1997
A:Description: The sequence of *C. elegans* cosmid C07G3.
A:Reference number: Z21080
A:Accession: T31758
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: DNA
A:Residues: 1-351 <GEI>
A:Cross-references: UNIPROT:O16326; EMBL:AF016432; PIDN:AAB65377.1; GSPDB:GN000023; CESP:
A:Experimental source: strain Bristol N2; clone C07G3
C:Genetics:
A:Gene: CESP:C07G3.3
A:Map position: 5
A:Introns: 66/3; 122/2; 167/2; 214/3; 319/1

Query Match 63.6%; Score 42; DB 2; Length 351;
Best Local Similarity 100.0%; Pred. No. 47;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CYWGC 5
|||||
Db 97 CYWGC 101

RESULT 9

I54210
N-acetylglactosamine-4-sulfatase (EC 3.1.6.12) precursor - rat (fragment)
N:Alternate names: arylsulfatase (EC 3.1.6.1) B [misidentification]
C:Species: *Rattus norvegicus* (Norway rat)
C>Date: 12-Aug-1996 #sequence_revision 12-Aug-1996 #text_change 09-Jul-2004
C:Accession: I54210

R:Kunieda, T.
Genomics 29, 582-587, 1995
A:Title: Mucopolysaccharidosis type VI in rats: Isolation of cDNAs encoding arylsulfatase
A:Reference number: I54210; MUID:96121368; PMID:8575749
A:Accession: I54210
A:Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-473 <RES>
A:Cross-references: UNIPROT:P50430; GB:D49434; NID:g1065603; PIDN:BAA08412.1; PID:g1089
C:Genetics:
A:Gene: ARSB
C:Superfamily: animal sulfatase
C:Keywords: sulfuric ester hydrolase
F:31/Modified site: 3-oxoalanine (Cys) #status predicted
Query Match 63.6%; Score 42; DB 2; Length 473;
Best Local Similarity 85.7%; Pred. No. 59;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 2 YMGCGYW 8
|||||
Db 384 YFGCGYW 390
RESULT 10
KJHUAB
N-acetylglactosamine-4-sulfatase (EC 3.1.6.12) precursor [validated] - human
N:Alternate names: arylsulfatase B (ASB); chondroitinase; chondroitinsulfatase; G4S; N-
C:Species: *Homo sapiens* (man)
C>Date: 31-Dec-1993 #sequence_revision 27-Oct-1995 #text_change 09-Jul-2004
C:Accession: S35990; S33307; A35078; A45659; A42449; B42449; C42449; I54217; A56865
R:Peters, C.W.B.
submitted to the EMBL data Library, March 1993
A:Reference number: S35990
A:Accession: S35990
A:Molecule type: DNA
A:Residues: 1-533 <PET>
A:Cross-references: UNIPROT:P15848; EMBL:X72735; NID:g289009; PIDN:CAA51272.1; PID:g825
R:Modaresi, S.; Rupp, K.; von Figura, K.; Peters, C.
Biol. Chem. Hoppe-Sevler 374, 327-335, 1993
A:Title: Structure of the human arylsulfatase B gene.
A:Reference number: S33307; MUID:93332648; PMID:7687847
A:Accession: S33307
A:Molecule type: DNA
A:Residues: 1-104 <MOD>
A:Cross-references: EMBL:X72735; EMBL:X72736; EMBL:X72737; EMBL:X72738; EMBL:X72739; EM
A:Note: the enzyme is referred to as EC 3.1.6.9
R:Peters, C.; Schmidt, B.; Rommerskirch, W.; Rupp, K.; Zuehlendorf, M.; Vingron, M.; Mey
J. Biol. Chem. 265, 3374-3381, 1990
A:Title: Phylogenetic conservation of arylsulfatases. cDNA cloning and expression of hu
A:Reference number: A35078; MUID:90153994; PMID:2303452
A:Accession: A35078
A:Molecule type: mRNA
A:Residues: 1-357, 'V', 359-533 <PE2>
A:Cross-references: GB:J05225; NID:g179076; PIDN:AAA51784.1; PID:g179077
A:Note: parts of this sequence were determined by protein sequencing
R:Litjens, T.; Morris, C.P.; Gibson, G.J.; Beckmann, K.R.; Hopwood, J.J.
Biochem. Int. 24, 209-215, 1991
A:Title: Human N-acetylglactosamine-4-sulphatase: protein maturation and isolation of
A:Reference number: A45659; MUID:92028992; PMID:1930244
A:Accession: A45659
A:Molecule type: DNA; protein
A:Residues: 1-104 <LIT>
A:Cross-references: GB:IS5777; NID:g236697; PIDN:AAB19988.1; PID:g236698
A:Note: sequence extracted from NCBI backbone (NCBIN:57777, NCSIP:57778)
A:Note: the enzyme is referred to as EC 3.1.6.1
A:Note: parts of this sequence, including the amino end of the mature protein, were det.
A:Note: a form is described with a proteolytic cleavage somewhere between residue 450 and
R:Jin, W.D.; Jackson, C.E.; Desnick, R.J.; Schuchman, E.H.
Am. J. Hum. Genet. 50, 793-800, 1992
A:Title: Mucopolysaccharidosis type VI: identification of three mutations in the arylsu
eity.

A:Reference number: A42449; MUID:92197625; PMID:1550123
A:Accession: A42449
A:Molecule type: mRNA
A:Residues: 115-116,'R',118 <JIN>
A:Cross-references: GB:S90729; NID:g247486; PIDN:AAB21831.1; PID:g247487
A>Note: sequence extracted from NCBI backbone (NCBIN:90729, NCBI:P:90731)
A:Accession: B42449
A:Molecule type: mRNA
A:Residues: 234-235,'P',237-238 <J12>
A:Cross-references: GB:S90736; NID:g247488; PIDN:AAB21832.1; PID:g247489
A>Note: sequence extracted from NCBI backbone (NCBIN:90736, NCBI:P:90739)
A:Accession: C42449
A:Molecule type: mRNA
A:Residues: 403-404,'Y',406-407 <J13>
A:Cross-references: GB:S90743; NID:g247490; PIDN:AAB21833.1; PID:g247491
A>Note: sequence extracted from NCBI backbone (NCBIN:90743, NCBI:P:90747)
A>Note: the enzyme is referred to as EC 3.1.6.1
A>Note: these mutations give rise to mucopolysaccharidosis type VI, Maroteaux-Lamy disease
R:Schmidt, B.; Selmer, T.; Ingendoh, A.; von Figura, K.
Cell 82, 271-278, 1995
A:Title: A novel amino acid modification in sulfatases that is defective in multiple sul
A:Reference number: AS7113; MUID:95354208; PMID:7628016
A:Contents: annotation; identification of 3-oxoalanine, 2-amino-3-oxopropanoic acid
R:Schuchman, E.H.; Jackson, C.E.; Deenick, R.J.
Genomics 6, 149-158, 1990
A:Title: Human arylsulfatase B: MOPAC cloning, nucleotide sequence of a full-length cDNA
A:Reference number: I54217; MUID:90152677; PMID:1968043
A:Accession: I54217
A>Status: preliminary; translated from GB/EMBL/DBJ
A:Molecule type: mRNA
A:Residues: 1-357,'V',359-375,'M',377-533 <RES>
A:Cross-references: GB:M32373; NID:gl79029; PIDN:AAA51779.1; PID:gl79030
R:Kobayashi, T.; Honke, K.; Jin, T.; Gasa, S.; Miyazaki, T.; Makita, A.
Biochim. Biophys. Acta 1159, 243-247, 1992
A:Title: Components and proteolytic processing sites of arylsulfatase B from human plac
A:Reference number: A56865; MUID:93003385; PMID:1390929
A:Accession: A56865
A:Molecule type: protein
A:Residues: 41-55;424-425,'X',427-454;466-483 <KOB>
A:Experimental source: placenta
A>Note: sequence modified after extraction from NCBI backbone
A>Note: the fragments shown are the amino ends of the alpha, gamma, and beta chains of t
C:Comment: This enzyme is frequently misidentified as EC 3.1.6.1.
C:Genetics:
A:Gene: GDB:ARSB
A:Cross-references: GDB:119008; OMIM:253200
A:Map position: 5q11-5q13
A:Introns: 104/3; 167/1; 230/3; 300/1; 381/2; 405/1; 446/1
A>Note: defects in this gene can cause mucopolysaccharidosis type VI, Maroteaux-Lamy dis
C:Function:
A:Description: hydrolyzes N-acetylgalactosamine-4-sulfate units in chondroitin sulfate a
C:Superfamily: animal sulfatase
C:Keywords: glycoprotein; lysosomal storage disease; lysosome; sulfuric ester hydrolase
F:1-40/Domain: signal sequence #status predicted <SIG>
F:41-423/Product: alpha chain #status predicted <AMAT>
F:424-465/Product: gamma chain #status predicted <GMAT>
F:466-533/Product: beta chain #status experimental <EMAT>
F:91/Modified site: 3-oxoalanine (Cys) #status experimental
F:188,279,366,458/Binding site: carboxylate (Asn) (covalent) #status predicted
F:291/Binding site: carboxylate (Asn) (covalent) #status absent
F:426/Binding site: carboxylate (Asn) (covalent) #status experimental

Query Match 63.6%; Score 42; DB 1; Length 533;
Best Local Similarity 85.7%; Pred. No. 65;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YWCGGYW 8
| | | | |
Db 444 YPCGGYW 450

RESULT 11
D70770

probable glycogen phosphorylase - Mycobacterium tuberculosis (strain H37RV)
C:Species: Mycobacterium tuberculosis
C:Date: 17-Jul-1998 #sequence_revision 17-Jul-1998 #text_change 09-Jul-2004
C:Accession: D70770
R:Cole, S.T.; Brosch, R.; Parkhill, J.; Garnier, T.; Churcher, C.; Harris, D.; Gordon, S.;
Connor, R.; Davies, R.; Devlin, K.; Feltwell, T.; Gentles, S.; Hamlin, N.; Holroyd, S.;
Rajandream, M.A.; Rogers, J.; Rutter, S.; Seeger, K.; Skelton, S.; Squares, S.
Nature 393, 537-544, 1998
A:Authors: Sgares, R.; Sulston, J.E.; Taylor, K.; Whitehead, S.; Barrell, B.G.
A:Title: Deciphering the biology of Mycobacterium tuberculosis from the complete genome
A:Reference number: A70500; MUID:98295987; PMID:9634230
A:Accession: D70770
A>Status: preliminary; nucleic acid sequence not shown; translation not shown
A:Molecule type: DNA
A:Residues: 1-863 <COI>
A:Cross-references: UNIPROT:Q10839; GB:Z73902; GB:AL123456; NID:g3261576; PIDN:CAA98092
A:Experimental source: strain H37RV
C:Genetics:
A:Gene: glgP
C:Superfamily: glucan phosphorylase

Query Match 63.6%; Score 42; DB 2; Length 863;
Best Local Similarity 71.4%; Pred. No. 95;
Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 2 YWCGGYW 8
| | | | |
Db 595 YWCGDWW 601

RESULT 12
B72392
hypothetical protein - Thermotoga maritima (strain MSB8)
C:Species: Thermotoga maritima
C:Date: 11-Jun-1999 #sequence_revision 11-Jun-1999 #text_change 09-Jul-2004
C:Accession: B72392
R:Nelson, K.E.; Clayton, R.A.; Gill, S.R.; Gwinn, M.L.; Dodson, R.J.; Haft, D.H.; Hicke
Garrett, M.M.; Stewart, A.M.; Cotton, M.D.; Pratt, M.S.; Phillips, C.A.; Richardson, D.;
C.M.
Nature 399, 323-329, 1999
A:Title: Evidence for lateral gene transfer between Archaea and Bacteria from genome seq
A:Reference number: A72200; MUID:99287316; PMID:10360571
A:Accession: B72392
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-83 <ARN>
A:Cross-references: UNIPROT:Q9WYF1; GB:AE001713; GB:AE000512; NID:g4980809; PIDN:AAD354
A:Experimental source: strain MSB8
C:Genetics:
A:Gene: TM0315

Query Match 62.1%; Score 41; DB 2; Length 83;
Best Local Similarity 50.0%; Pred. No. 21;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY 1 CVWGGGYW 8
| | | | |
Db 5 CFWSWGF 12

RESULT 13
G84839
late embryogenesis abundant M17 protein [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C:Date: 02-Feb-2001 #sequence_revision 02-Feb-2001 #text_change 09-Jul-2004
C:Accession: G84839
R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.;
M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.
euss, D.; Niernan, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter,
Nature 402, 761-768, 1999
A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.
A:Reference number: A84420; MUID:20083487; PMID:10617197
A:Accession: G84839

A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-280 <STO>
A;Cross-references: UNIPROT:Q9S783; GB:AE002093; NID:g3894196; PIDN:AACT78545.1; GSPDB:GN
C;Genetics:
A;Gene: At2g41260
A;Map position: 2

Query Match 62.1%; Score 41; DB 2; Length 280;
Best Local Similarity 60.0%; Pred. No. 55;
Matches 6; Conservative 1; Mismatches 1; Indels 2; Gaps 1;

QY 1 CYWGC--GYW 8
| | | | |
DB 78 CRWCCCGGW 87

RESULT 14
H96968
Integral membrane protein similar to antibiotic resistance protein B. subtilis [imported]
C;Species: Clostridium acetobutylicum
C;Date: 14-Sep-2001 #sequence_revision 14-Sep-2001 #text_change 09-Jul-2004
C;Accession: H96968
.; Nolling, J.; Bretton, G.; Omelchenko, M.V.; Markarova, K.S.; Zeng, Q.; Gibson, R.; Lee,
.; Daly, M.J.; Bennett, G.N.; Koonin, E.V.; Smith, D.R.
J. Bacteriol. 183, 4823-4838, 2001
A;Title: Genome Sequence and Comparative Analysis of the Solvent-Producing Bacterium Clo
A;Reference number: A96900; MUID:21359325; PMID:21359325
A;Accession: H96968
A;Status: preliminary
A;Molecule type: DNA
A;Residues: 1-441 <KUR>
A;Cross-references: UNIPROT:Q97LK0; GB:AE001437; PIDN:AAK78539.1; PID:g15023427; GSPDB:G
A;Experimental source: Clostridium acetobutylicum ATCC824
C;Genetics:
A;Gene: CAC0560

Query Match 62.1%; Score 41; DB 2; Length 441;
Best Local Similarity 71.4%; Pred. No. 78;
Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CYWGCY 7
| | | | |
DB 333 CFWGMY 339

RESULT 15
A29476
muscarinic acetylcholine receptor M4 - rat
C;Species: Rattus norvegicus (Norway rat)
C;Date: 31-Dec-1988 #sequence_revision 31-Dec-1988 #text_change 09-Jul-2004
C;Accession: A29476
R.; Braun, T.; Schofield, P.R.; Shivers, B.D.; Pritchett, D.B.; Seeburg, P.H.
Biochem. Biophys. Res. Commun. 149, 125-132, 1987
A;Title: A novel subtype of muscarinic receptor identified by homology screening.
A;Reference number: A29476; MUID:88077068; PMID:3120722
A;Accession: A29476
A;Molecule type: mRNA
A;Residues: 1-589 <BRA>
A;Cross-references: UNIPROT:P08483; GB:M18088; NID:g202657; PIDN:AAA40659.1; PID:g202658
A;Experimental source: brain
C;Superfamily: vertebrate rhodopsin
C;Keywords: G protein-coupled receptor; glycoprotein; neurotransmitter receptor; transme
F;67-90/Domain: transmembrane #status predicted <TM1>
F;104-124/Domain: transmembrane #status predicted <TM2>
F;142-163/Domain: transmembrane #status predicted <TM3>
F;184-206/Domain: transmembrane #status predicted <TM4>
F;230-251/Domain: transmembrane #status predicted <TM5>
F;492-512/Domain: transmembrane #status predicted <TM6>
F;527-545/Domain: transmembrane #status predicted <TM7>

Query Match 62.1%; Score 41; DB 2; Length 589;
Best Local Similarity 71.4%; Pred. No. 98;

Matches 5; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 2 YWCCGYW 8
| | | | |
DB 524 YWNLGYW 530

Search completed: April 18, 2005, 19:52:01
Job time : 19.4419 secs

1115 Page Blank (uspio)

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 19:37:28 ; Search time 71.2558 Seconds
(without alignments)
57.492 Million cell updates/sec

Title: SEQ5
Perfect score: 66
Sequence: 1 cywgcgyw 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1612378 seqs, 512079187 residues

Total number of hits satisfying chosen parameters: 1612378

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Uniprot_03.*

1: uniprot_sprot.*
2: uniprot_trembl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	72.7	1182	2 Q9ENL3	Q9enl3 colorado ti
2	47.5	72.0	134	2 Q7S2D0	Q7s2d0 neurospora
3	47	71.2	841	2 Q7MU18	Q7mul18 porphyromon
4	46	69.7	181	2 Q728W6	Q728w6 desulfovibr
5	46	69.7	208	2 Q8KL42	Q8kl42 rhizobium e
6	46	69.7	518	1 SAP_CHICK	Ol3035 gallus gall
7	46	69.7	553	2 Q6P7A4	Q6p7a4 rattus norv
8	46	69.7	554	1 SAP_RAT	P10960 rattus norv
9	45.5	68.9	93	2 Q6VY22	Q6vy22 oryza sativ
10	45.5	68.9	269	2 Q9C9X6	Q9c9x6 arabidopsis
11	45.5	68.9	279	2 Q9LDE2	Q9lde2 arabidopsis
12	45.5	68.9	283	2 Q6UIZ7	Q6uiz7 arabidopsis
13	45.5	68.9	283	2 Q9SKD0	Q9skd0 arabidopsis
14	45.5	68.9	285	2 Q8L9F3	Q8l9f3 arabidopsis
15	45	68.2	463	2 Q7MYM7	Q7mym7 photorhabdu
16	44	66.7	840	2 Q8A6N9	Q8a6n9 bacteroides
17	44	66.7	845	2 Q6QOS3	Q6qos3 bacteroides
18	43	65.2	78	2 Q6IE40	Q6ie40 rattus norv
19	43	65.2	135	2 Q6DTM1	Q6dtm1 meleagrid h
20	43	65.2	145	2 Q6DTM2	Q6dtm2 meleagrid h
21	43	65.2	161	2 Q6LU09	Q6lu09 photobacter
22	43	65.2	239	2 Q69T79	Q69t79 oryza sativ
23	42	63.6	64	2 Q9UD19	Q9ud19 homo sapien
24	42	63.6	182	1 RELX_HORSE	P22969 equus cabal
25	42	63.6	185	2 Q6AZ36	Q6az36 rattus norv
26	42	63.6	287	2 Q96KT3	Q96kt3 homo sapien
27	42	63.6	314	2 Q8A188	Q8a188 bacteroides
28	42	63.6	332	1 CGFI_HUMAN	Q98j75 homo sapien
29	42	63.6	332	1 CGFI_MOUSE	Q98j77 mus musculu
30	42	63.6	332	1 CGFI_RAT	P97587 rattus norv
31	42	63.6	342	2 Q16325	Q16325 caenorhabdi

32	42	63.6	347	2	Q16323	Q16323 caenorhabdi
33	42	63.6	351	2	Q16326	Q16326 caenorhabdi
34	42	63.6	473	1	ARSB_RAT	P50430 rattus norv
35	42	63.6	533	1	ARSB_HUMAN	P15848 homo sapien
36	42	63.6	863	1	PHSG_WYCTU	Q10639 mycobacteri
37	42	63.6	863	2	Q7U078	Q7u078 mycobacteri
38	42	63.6	871	2	Q73X77	Q73x77 mycobacteri
39	42	63.6	2954	2	Q96898	Q96898 hepatitis g
40	42	63.6	2967	2	Q41892	Q41892 hepatitis g
41	41	62.1	83	2	Q9WVF1	Q9wvf1 thermotoga
42	41	62.1	123	2	Q65XY5	Q65xy5 caenorhabdi
43	41	62.1	126	2	Q9WNF1	Q9wnf1 drosophila
44	41	62.1	225	2	Q8L7T1	Q8l7t1 arabidopsis
45	41	62.1	252	2	P95462	P95462 plectonema

ALIGNMENTS

RESULT 1

Q9ENL3 PRELIMINARY; PRT; 1182 AA.
AC Q9ENL3;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-DEC-2001 (Tremblrel. 19, Last annotation update)
DE VP3.
OS Colorado tick fever virus.
OC Viruses; dsRNA viruses; Reoviridae; Coltivirus.
OX NCBI_TaxID=46839;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=Florio;
RC MEDLINE=20351236; PubMed=10891382; DOI=10.1006/bbrc.2000.3057;
RA Attoui H., Billoir F., Biagini P., Cantaloube J.F., de Chesse R.,
De Micco P., de Lamballerie X.;
RT "Sequence determination and analysis of the full-length genome of
RT colorado tick fever virus, the type species of genus Coltivirus
RT (Family Reoviridae).";
RL Biochem. Biophys. Res. Commun. 273:1121-1125(2000).
DR EMBL; AF139759; FAG00068.1; -;
SQ SEQUENCE 1182 AA; 135010 MW; E1F5F3515E9204C9 CRC64;

Query Match 72.7%; Score 48; DB 2; Length 1182;
Best Local Similarity 100.0%; Pred. No. 72;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CYWGC 6
|||
Db 480 CYWGC 485

RESULT 2

Q7S2D0 PRELIMINARY; PRT; 134 AA.
AC Q7S2D0;
DT 01-MAR-2004 (Tremblrel. 26, Created)
DT 01-MAR-2004 (Tremblrel. 26, Last sequence update)
DT 01-MAR-2004 (Tremblrel. 26, Last annotation update)
DE Predicted protein.
GN Name=NCU05948.1;
OS Neurospora crassa.
OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;
OC Sordariomycetidae; Sordariales; Sordariaceae; Neurospora.
OX NCBI_TaxID=5141;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=OR74A;
RA Galagan J.E., Calvo S.E., Borkovich K.A., Selker E.U., Read N.D.,
Jaffe D., Fitzhugh W., Ma L.-J., Smirnov S., Purcell S., Rehman B.,
Elkins T., Engels R., Wang S., Nielsen C.B., Butler J., Endrizzi M.,
Qui D., Iannakiev P., Pedersen D., Nelson M., Washburne M.,
Selitrennikoff C.P., Kinsey J.A., Braun E.L., Zelter A., Schulte U.,

```

RA Kothe G.O., Jedd G., Mewes W., Staben C., Marcotte E., Greenberg D.,
RA Roy A., Foley K., Navlyar J., Thomann N., Barrett R., Gnerre S.,
RA Kamal M., Kamvysseilis M., Mauceli E., Bielke C., Rudd S., Frishman D.,
RA Krystofova S., Rasmussen C., Metzenberg R.L., Perkins D.D., Kroken S.,
RA Cogoni C., Macino G., Catcheside D., Li W., Pratt R.J., Osmari S.A.,
RA DeSouza C.C., Glass L., Orbach M.J., Berglund J., Voelker R.,
RA Yarden O., Plamann M., Seiler S., Dunlap J., Radford A., Aramayo R.,
RA Natvig D.O., Alex L.A., Mannhaupt G., Ebbola D.J., Freitag M.,
RA Paulsen I., Sachs M.S., Lander E.S., Nusbaum C., Birren B.,
RT "The Genome Sequence of the Filamentous Fungus Neurospora crassa.";
RL Nature 0:0-0(2003).
CC -!- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; AABX01000427; EAA29556.1; -.
SQ SEQUENCE 134 AA; 15465 MW; 39231A0E10C7D57B CRC64;

Query Match 72.0%; Score 47.5; DB 2; Length 134;
Best Local Similarity 53.8%; Pred. No. 14;
Matches 7; Conservative 0; Mismatches 1; Indels 5; Gaps 1;

QY 1 CY-----WCGGYW 8
Db 2 CYFEOTLWSCGYW 14

RESULT 3
Q7MUL8 PRELIMINARY; PRT; 841 AA.
AC Q7MUL8;
DT 01-MAR-2004 (TrEMBLrel. 26, Created)
DT 01-WAR-2004 (TrEMBLrel. 26, Last sequence update)
DT 01-WAR-2004 (TrEMBLrel. 26, Last annotation update)
DE Conserved domain protein.
GN OrderedLocNames=PGI754;
OS Porphyromonas gingivalis (Bacteroides gingivalis).
OC Bacteria; Bacteroidetes; Bacteroides (class); Bacteroidales;
OC Porphyromonadaceae; Porphyromonas.
OX NCBI_TaxID=837;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=W83;
RX DOI=10.1128/JB.185.18.5591-5601.2003;
RA Nelson K.E., Fleischmann R.D., DeBoy R.T., Paulsen I.T., Fouts D.E.,
RA Eisen J.A., Daugherty S.C., Dodson R.J., Durkin A.S., Gwinn M.L.,
RA Haft D.H., Kolonay J.F., Nelson W.C., Mason T.M., Tallon L., Gray J.,
RA Granger D., Tettelin H., Dong H., Galvin J.L., Duncan M.J.,
RA Dewhirst F.E., Fraser C.M.;
RT "Complete genome sequence of the oral pathogenic bacterium
RT Porphyromonas gingivalis strain W83."
RL J. Bacteriol. 185:5591-5601(2003).
DR EMBL; AEO17178; AAQ66756.1; -.
DR TIGR; PG1754; -.
DR GO; GO:0003824; F: catalytic activity; IEA.
DR InterPro; IPR001375; Peptidase_S9.
DR InterPro; IPR011047; Quin_alc_DH_like.
DR InterPro; IPR000379; Ser_estr.
DR Pfam; PF00326; Peptidase_S9; 1.
KW Complete proteome.
SQ SEQUENCE 841 AA; 94237 MW; A082DCD732EC3F36 CRC64;

Query Match 71.2%; Score 47; DB 2; Length 841;
Best Local Similarity 85.7%; Pred. No. 76;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YWCGGYW 8
Db 719 YWGSQYW 725

RESULT 4
Q728W6

```

```

ID Q728W6 PRELIMINARY; PRT; 181 AA.
AC Q728W6;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Acetyltransferase, GNAT family.
GN OrderedLocNames=DVU2486;
OS Desulfovibrio vulgaris (strain Hildenborough / ATCC 29579 / NCIMB
OS 8303).
OC Bacteria; Proteobacteria; Deltaproteobacteria; Desulfovibrionales;
OC Desulfovibrionaceae; Desulfovibrio.
OX NCBI_TaxID=882;
RN [1]
RP SEQUENCE FROM N.A.
RX PubMed=15077118; DOI=10.1038/nbt959;
RA Heidelberg J.F., Seshadri R., Haveman S.A., Hemme C.L., Paulsen I.T.,
RA Kolonay J.F., Eisen J.A., Ward N.L., Methe B.A., Brinkac L.M.,
RA Daugherty S.C., DeBoy R.T., Dodson R.J., Durkin A.S., Madupu R.,
RA Nelson W.C., Sullivan S.A., Fouts D.E., Haft D.H., Selengut J.,
RA Peterson J.D., Davidson T.M., Zafar N., Zhou L., Radune D.,
RA Dinitrov G., Hance M., Tran K., Khouri H.M., Gill J., Utterback T.R.,
RA Feldblyum T.V., Wall J.D., Voordouw G., Fraser C.M.;
RT "The genome sequence of the anaerobic, sulfate-reducing bacterium
RT Desulfovibrio vulgaris Hildenborough."
RL Nat. Biotechnol. 22:554-559(2004).
DR EMBL; AEO17317; AAS96958.1; -.
DR TIGR; DVU2486; -.
DR GO; GO:0008080; F: N-acetyltransferase activity; IEA.
DR GO; GO:0016740; F: transferase activity; IEA.
DR InterPro; IPR000182; GCN5acetyl_trans.
DR Pfam; PF00583; Acetyltransf_1; 1.
KW Complete proteome; Transferase.
SQ SEQUENCE 181 AA; 20659 MW; 71621F87EDC147D9 CRC64;

Query Match 69.7%; Score 46; DB 2; Length 181;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 6; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 2 YWCGGY 7
Db 103 YWCGGY 108

RESULT 5
Q8KL42 PRELIMINARY; PRT; 208 AA.
ID Q8KL42
AC Q8KL42;
DT 01-OCT-2002 (TrEMBLrel. 22, Created)
DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE Hypothetical protein yh031.
GN Name=yh031;
OS Rhizobium etli.
OG Plasmid symbiotic plasmid p42d.
OC Bacteria; Proteobacteria; Alphaproteobacteria; Rhizobiales;
OC Rhizobiaceae; Rhizobium/Agrobacterium group; Rhizobium.
OX NCBI_TaxID=29449;
RN [1]
RP SEQUENCE FROM N.A.
RX STRAIN=CFN42;
RC MEDLINE=91193195; PubMed=2013564;
RA Girard M.L., Flores M., Brom S., Romero D., Palacios R., Davila G.;
RT "Structural complexity of the symbiotic plasmid of Rhizobium
RT leguminosarum bv. phaseoli."
RL J. Bacteriol. 173:2411-2419(1991).
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN42;
RX MEDLINE=97419521; PubMed=9274036;
RA Ramirez-Romero M.A., Bustos P., Girard L., Rodriguez O.,
RA Cevallos M.A., Davila G.;
RT "Sequence, localization and characteristics of the replicator region
RT of the symbiotic plasmid of Rhizobium etli."

```

```

RL Microbiology 143:2825-2831(1997).
RN [3]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN42;
RX MEDLINE=22309397; PubMed=12421308;
RA Quintero V., Cevallos M.A., Davila G.;
RT "A site-specific recombinase (RinQ) is required to exert
RL incompatibility towards the symbiotic plasmid of Rhizobium etli.";
RN Mol. Microbiol. 46:1023-1032(2002).
RN [4]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN42;
RA Ramirez M.A., Bustos P., Girard L., Rodriguez O., Cevallos M.A.,
RA Davila G.;
RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
RN [5]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN42;
RA Quintero V., Bustos P., Davila G.;
RL Submitted (MAY-2002) to the EMBL/GenBank/DBJ databases.
RN [6]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN42;
RA Gonzalez V., Bustos P., Medrano-Soto A., Ramirez-Romero M.A.,
RA Romero D., Salgado H., Hernandez-Gonzalez I., Hernandez-Celis J.C.,
RA Quintero V., Girard L., Rodriguez O., Flores M., Cevallos M.A.,
RA Collado-Vides J., Davilla G.;
RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
RN [7]
RP SEQUENCE FROM N.A.
RC STRAIN=CFN42;
RA Quintero V., Bustos P., Davila G.;
RL Submitted (JUN-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; U80928; AAMS4925.1; -
KW Hypothetical protein; Plasmid.
SQ SEQUENCE 208 AA; 23594 MW; 2DF473380800744 CRC64;

Query Match 69.7%; Score 46; DB 2; Length 208;
Best Local Similarity 75.0%; Pred. No. 33;
Matches 6; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 1 CWFGCGYW 8
DB 7 CYFGSGYW 14

RESULT 6
SAP_CHICK STANDARD; PRT; 518 AA.
AC O13035;
DT 16-OCT-2001 (Rel. 40, Created)
DT 16-OCT-2001 (Rel. 40, Last sequence update)
DT 05-JUL-2004 (Rel. 44, Last annotation update)
DE Proactivator polypeptide precursor [Contains: Saposin A; Saposin B;
DE Saposin C; Saposin D].
GN Name=PGAP;
OS Gallus Gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP SEQUENCE FROM N.A., AND SEQUENCE OF 194-203.
RC TISSUE=Brain, and Liver;
RX MEDLINE=98129745; PubMed=9461526;
RA Azuma N., Seo H.-C., Lie O., Fu Q., Gould R.M., Hiraiwa M., Burt D.W.,
RA Paton I.R., Morrice D.R., O'Brien J.S., Kishimoto Y.;
RT "Cloning, expression and map assignment of chicken prosaposin.";
RL Biochem. J. 330:321-327(1998).
RN [2]
RP SEQUENCE FROM N.A.
RA Altman N., Horowitz M.;
RL Submitted (NOV-1998) to the EMBL/GenBank/DBJ databases.

-!- FUNCTION: The lysosomal degradation of sphingolipids takes place
CC by the sequential action of specific hydrolases. Some of these
CC enzymes require specific low-molecular mass, non-enzymic proteins:
CC the sphingolipids activator proteins (coproteins) (By similarity).
CC -!- FUNCTION: Saposin A and saposin C stimulate the hydrolysis of
CC glucosylceramide by beta-glucosylceramidase (EC 3.2.1.45) and
CC galactosylceramide by beta-galactosylceramidase (EC 3.2.1.46).
CC Saposin C apparently acts by combining with the enzyme and acidic
CC lipid to form an activated complex, rather than by solubilizing
CC the substrate (By similarity).
CC -!- FUNCTION: Saposin B stimulates the hydrolysis of galacto-
CC cerebroside sulfate by arylsulfatase A (EC 3.1.6.8), GM1
CC gangliosides by beta-galactosidase (EC 3.2.1.23) and
CC globotriaosylceramide by alpha-galactosidase A (EC 3.2.1.22).
CC Saposin B forms a solubilizing complex with the substrates of the
CC sphingolipid hydrolases (By similarity).
CC -!- FUNCTION: Saposin D is a specific sphingomyelin phosphodiesterase
CC activator (EC 3.1.4.12) (By similarity).
CC -!- SUBUNIT: Saposin B is a homodimer (By similarity).
CC -!- SUBCELLULAR LOCATION: Lysosomal (By similarity).
CC -!- PTM: This precursor is proteolytically processed to 4 small
CC peptides, which are similar to each other and are sphingolipid
CC hydrolase activator proteins (By similarity).
CC -!- SIMILARITY: Contains 2 saposin A-type domains.
CC -!- SIMILARITY: Contains 4 saposin B-type domains.
-----
This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
-----
DR EMBL; AB003471; BAA19914.1; -
DR EMBL; AF108656; AAF05899.1; -
DR HSSP; Q92739; IN69.
DR InterPro; IPR003119; SapaA.
DR InterPro; IPR007856; SapaB_1.
DR InterPro; IPR008138; SapaB_2.
DR InterPro; IPR008140; SapaB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR011001; Saposin like.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; Sapa; 2.
DR Pfam; PF05184; SapaB_1; 4.
DR Pfam; PF03489; SapaB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapaB_sub; 1.
DR SMART; SM00162; SAPA; 2.
DR SMART; SM00741; SapaB; 4.
DR Direct protein sequencing; Glycoprotein; GM2-gangliosidosis; Lysosome;
KW Repeat; Signal; Sphingolipid metabolism.
FT SIGNAL 1 17
FT PROPEP 18 60
FT CHAIN 61 143 Saposin A.
FT PROPEP 145 193
FT CHAIN 194 276 Saposin B.
FT PROPEP 278 305 Saposin C.
FT CHAIN 307 387 Saposin D.
FT PROPEP 389 398
FT CHAIN 399 480
FT PROPEP 482 518
FT DOMAIN 22 55 Saposin-like type A 1.
FT DOMAIN 60 143 Saposin-like type B 1.
FT DOMAIN 193 277 Saposin-like type B 2.
FT DOMAIN 307 388 Saposin-like type B 3.
FT DOMAIN 399 480 Saposin-like type B 4.
FT DOMAIN 485 518 Saposin-like type A 2.
FT DISULFID 64 139 By similarity.
FT DISULFID 67 133 By similarity.
FT DISULFID 95 107 By similarity.
FT DISULFID 197 273 By similarity.

```

```

FT DISULFID 200 267 By similarity.
FT DISULFID 229 240 By similarity.
FT DISULFID 311 384 By similarity.
FT DISULFID 314 378 By similarity.
FT DISULFID 342 353 By similarity.
FT DISULFID 403 476 By similarity.
FT DISULFID 406 470 By similarity.
FT DISULFID 434 445 By similarity.
FT CARBOHYD 81 81 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 214 214 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 328 328 N-linked (GlcNAc...) (Potential).
FT CARBOHYD 420 420 N-linked (GlcNAc...) (Potential).
FT CONFLICT 94 94 R -> T (in Ref. 2).
FT CONFLICT 486 486 E -> D (in Ref. 2).
SQ SEQUENCE 518 AA; 57601 MW; B803000E991C3963 CRC64;

Query Match 69.7%; Score 46; DB 1; Length 518;
Best Local Similarity 75.0%; Pred. No. 71;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CYWGGGYW 8
DB 489 CVMGPGYW 496

RESULT 7
QSP7A4 PRELIMINARY; PRT; 553 AA.
AC Q6P7A4;
DT 05-JUL-2004 (TRENBLrel. 27, Created)
DE Prosaposin.
GN Name=Psap;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hong L.,
RA Diatchenko L., Maruina K., Farmer A.A., Rubin G.M., Hsieh F.,
RA Brownstein M.J., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahy J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., Butterfield Y.S.,
RA Krzywinski M.I., Skalska U., Smalios D.E., Schnerch A., Schein J.E.,
RA Jones S.J., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
and mouse cDNA sequences."
RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Prostate;
RA Strausberg R.;
RL Submitted (NOV-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC061759; AAH61759.1;
DR GO: GO:0005764; C:lysosome; IEA.
DR GO: GO:0006623; P:lipid metabolism; IEA.
DR GO: GO:0006665; P:sphingolipid metabolism; IEA.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; Sapa_1.

```

```

DR InterPro; IPR008138; SapaB_2.
DR InterPro; IPR008140; SapaB_sub.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR008139; SaposinB.
DR InterPro; IPR011001; Saposin_like.
DR Pfam; PF02199; Sapa; 2.
DR Pfam; PF05184; SapaB_1; 3.
DR Pfam; PF03489; SapaB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR PRODOM; PD001732; SapaB_sub; 1.
DR SMART; SMO0162; SAPA; 2.
SQ SEQUENCE 553 AA; 61039 MW; 63F3DD5E0C523393 CRC64;

Query Match 69.7%; Score 46; DB 2; Length 553;
Best Local Similarity 75.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CYWGGGYW 8
DB 524 CVMGPGYW 531

RESULT 8
SAP_RAT STANDARD; PRT; 554 AA.
ID SAP_RAT
AC P10560; O62841; O64190;
DT 01-JUL-1989 (Rel. 11, Created)
DT 01-JUL-1989 (Rel. 11, Last sequence update)
DE Sulfated glycoprotein 1 precursor (SGP-1) (Prosaposin).
GN Name=Psap; Synonyms=SGp1;
OS Rattus norvegicus (Rat).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Rattus.
OX NCBI_TaxID=10116;
RN [1]
RP SEQUENCE FROM N.A.; AND PARTIAL SEQUENCE.
RC TISSUE=Sertoli cells;
RX MEDLINE=89000647; PubMed=3048385;
RA Collard M.W., Sylvester S.R., Tsuruta J.K., Griswold M.D.;
RT "Biosynthesis and molecular cloning of sulfated glycoprotein 1
secreted by rat Sertoli cells: sequence similarity with the 70-
kilodalton precursor to sulfatide/GM1 activator."
RL Biochemistry 27:4557-4564(1988).
RN [2]
RP SEQUENCE FROM N.A.
RC TISSUE=Testis;
RX MEDLINE=96128541; PubMed=8573994;
RA Morales C.R., El-Alfy M., Zhao Q., Igldoura S.A.;
RT "Molecular role of sulfated glycoprotein-1 (SGP-1/prosaposin) in
Sertoli cells."
RL Histol. Histopathol. 10:1023-1034(1995).
RN [3]
RP SEQUENCE FROM N.A.
RC TISSUE=Testicle;
RX MEDLINE=96175245; PubMed=8601692;
RA Morales C.R., El-Alfy M., Zhao Q., Igldoura S.A.;
RT "Expression and tissue distribution of rat sulfated glycoprotein-1
(prosaposin)."
RL J. Histochem. Cytochem. 44:327-337(1996).
CC -!- SUBCELLULAR LOCATION: Extracellular.
CC -!- SIMILARITY: Contains 2 saposin A-type domains.
CC -!- SIMILARITY: Contains 4 saposin B-type domains.
CC
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (See http://www.isb-sib.ch/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; M19936; AAA42136.1; -.

```

```
DR EMBL; S81353; AAB36042.2; -.
DR EMBL; S81373; AAB36233.2; -.
DR PIR; A28716; A28716.
DR HSP; Q92739; IN69.
DR RGD; 3423; Psap.
DR InterPro; IPR003119; Sapa.
DR InterPro; IPR007856; SapB_1.
DR InterPro; IPR008138; SapB_2.
DR InterPro; IPR008140; SapB_2.
DR InterPro; IPR008373; Saposin.
DR InterPro; IPR011001; Saposin-like.
DR InterPro; IPR008139; SaposinB.
DR Pfam; PF02199; Sapa; 2.
DR Pfam; PF05184; SapB_1; 3.
DR Pfam; PF03489; SapB_2; 4.
DR PRINTS; PR01797; SAPOSIN.
DR ProDom; PD001732; SapB_sub; 1.
KW Direct protein sequencing; Glycoprotein; Repeat; Signal.
FT SIGNAL 1 16
FT CHAIN 17 554 Sulfated glycoprotein 1.
FT DOMAIN 21 54 Saposin-like type A 1.
FT FT DOMAIN 61 138 Saposin-like type B 1.
FT FT DOMAIN 193 274 Saposin-like type B 2.
FT FT DOMAIN 310 391 Saposin-like type B 3.
FT FT DOMAIN 435 516 Saposin-like type B 4.
FT FT DOMAIN 521 554 Saposin-like type A 2.
FT FT DISULFID 63 138 By similarity.
FT FT DISULFID 66 132 By similarity.
FT FT DISULFID 94 106 By similarity.
FT FT DISULFID 197 270 By similarity.
FT FT DISULFID 200 264 By similarity.
FT FT DISULFID 229 240 By similarity.
FT FT DISULFID 314 387 By similarity.
FT FT DISULFID 317 381 By similarity.
FT FT DISULFID 345 356 By similarity.
FT FT DISULFID 439 512 By similarity.
FT FT DISULFID 442 506 By similarity.
FT FT DISULFID 470 481 By similarity.
FT FT CARBOHYD 80 80 N-linked (GLCNAC. . .) (Potential).
FT FT CARBOHYD 214 214 N-linked (GLCNAC. . .) (Potential).
FT FT CARBOHYD 331 331 N-linked (GLCNAC. . .) (Potential).
FT FT CARBOHYD 456 456 N-linked (GLCNAC. . .) (Potential).
FT FT CONFLICT 115 115 P -> L (in Ref. 2 and 3).
FT FT CONFLICT 299 299 D -> E (in Ref. 2).
FT FT CONFLICT 462 462 I -> V (in Ref. 3).
FT FT CONFLICT 527 527 W -> R (in Ref. 3).
FT FT CONFLICT 536 536 S -> M (in Ref. 3).
FT SEQUENCE 554 AA; 61123 MW; DFE3F3A3A0520C6B CRC64;
Query Match 69.7%; Score 46; DB 1; Length 554;
Best Local Similarity 75.0%; Pred. No. 75;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 1 CYWGGY 8
DB 525 CVWPGY 532
RESULT 9
Q6Y22 PRELIMINARY; PRT; 93 AA.
AC Q6Y22;
DT 05-JUL-2004 (TrEMBLrel. 27, Created)
DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein OSJNB0056122.26;
GN Name=OSJNB0056122.26;
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]

RP SEQUENCE FROM N.A.
RA Sasaki T., Matsumoto T., Katayose Y.;
RL Submitted (AUG-2002) to the EMBL/GenBank/DBJ databases.
RW EMBL; AP005644; BAD17478.1; -.
KW Hypothetical protein.
SQ SEQUENCE 93 AA; 9536 MW; A206432A3764546F CRC64;
Query Match 68.9%; Score 45.5; DB 2; Length 93;
Best Local Similarity 63.6%; Pred. No. 20;
Matches 7; Conservative 0; Mismatches 1; Indels 3; Gaps 1;
QY 1 CYWGC---GYW 8
DB 76 CRWGCCHRGY 86
RESULT 10
Q9C9X6 PRELIMINARY; PRT; 269 AA.
AC Q9C9X6;
DT 01-JUN-2001 (TrEMBLrel. 17, Created)
DT 01-JUN-2001 (TrEMBLrel. 17, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE Hypothetical protein T23K23.3 (Basic pentacysteine 3).
GN Name=T23K23.3; Synonyms=BPC3; ORFNames=At1g68120;
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Lin X., Kaul S., Town C.D., Benito M.-I., Creasy T.H., Haas B.J.,
RA Wu D., Maiti R., Ronning C.M., Koo H., Fujii C.Y., Utterback T.R.,
RA Barnstead M.E., Bowman C.L., White O., Nierman W.C., Fraser C.M.;
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Town C.D., Kaul S.;
RL Submitted (JAN-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RX PubMed=14731261; DOI=10.1046/j.1365-3113X.2003.01971.x;
RA Meister R.J., Williams L.A., Monfared M.M., Gallagher T.L.,
RA Kraft E.A., Nelson C.G., Gasser C.S.;
RT "Definition and interactions of a positive regulatory element of the
RT Arabidopsis INNER NO OUTER promoter.";
RL Plant J. 37:426-438(2004).
DR EMBL; AC012563; AAG52002.1; -.
DR EMBL; AY380569; AAR25823.1; -.
DR PIR; E96704; E96704.
DR InterPro; IPR010409; DUF1004.
DR Pfam; PF06217; DUF1004; 1.
KW Hypothetical protein.
SQ SEQUENCE 269 AA; 30392 MW; 2753AB3AD6063FE8 CRC64;
Query Match 68.9%; Score 45.5; DB 2; Length 269;
Best Local Similarity 77.8%; Pred. No. 48;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 CY-WGCGY 8
DB 190 CYRWGCGGW 198
RESULT 11
Q9LDE2 PRELIMINARY; PRT; 279 AA.
AC Q9LDE2;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
DE F10B6.5 (TSE21.17) (At1g14680/F10B6_22) (Basic pentacysteine 2).
RN [1]
```

GN Name=BPC2; ORFNames=At1g14685;
 OS Arabidopsis thaliana (Mouse-ear cress);
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Chao Q., Shinn P., Dunn P., Buehler E., Kahn S., Kim C., Walker M.,
 RA Williams S., Altafi H., Araujo R., Conn L., Conway A.B., Gonzalez A.,
 RA Hansen N.F., Huizer L., Kremenetska I., Lenz C., Li J., Liu S.,
 RA Luros S., Rowley D., Schwartz J., Toriumi M., Vysotska V., Yu G.,
 RA Davis R.W., Federspiel N.A., Theologis A., Ecker J.R.;
 RL Submitted (MAY-1999) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,
 RA Khan S., Kim C., Altafi H., Bei B., Chin C., Chioi J., Choi E.,
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,
 RA Theologis A., Ecker J.;
 RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
 RA Shinn P., Tambunga G., Altafi H., Bei B., Chin C., Chioi J., Choi E.,
 RA Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,
 RA Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,
 RA Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,
 RA Thaveri A., Toriumi M., Vaysberg M., Yu G., Federspiel N.A.,
 RA Theologis A., Ecker J.R.;
 RL Submitted (NOV-1999) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
 RA Shinn P., Altafi H., Bei B., Chin C., Chioi J., Choi E., Conn L.,
 RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
 RA Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
 RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
 RA Ecker J.;
 RL Submitted (APR-2000) to the EMBL/GenBank/DBJ databases.
 RN [5]
 RP SEQUENCE FROM N.A.
 RA Chao Q., Brooks S., Buehler E., Johnson-Hopson C., Khan S., Kim C.,
 RA Shinn P., Altafi H., Bei B., Chin C., Chioi J., Choi E., Conn L.,
 RA Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B., Lee J.,
 RA Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N., Nguyen M.,
 RA Palm C., Pham P., Sakano H., Schwartz J., Southwick A., Thaveri A.,
 RA Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N., Theologis A.,
 RA Ecker J.;
 RL Submitted (OCT-2000) to the EMBL/GenBank/DBJ databases.
 RN [6]
 RP SEQUENCE FROM N.A.
 RA Chao Q., Chen H., Kim C.J., Koesema E., Meyers M.C., Banh J.,
 RA Bowser L., Carninci P., Dale J.M., Goldsmith A.D., Hayashizaki Y.,
 RA Ishida J., Jiang P.X., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Liu S.X., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Pham P.K., Quach H.L., Sakurai T.,
 RA Satou M., Seki M., Southwick A., Tang C.C., Toriumi M., Yamada K.,
 RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.
 RN [7]
 RP SEQUENCE FROM N.A.
 RA Kim C.J., Chen H., Cheuk R., Meyers M.C., Shinn P., Banh J.,
 RA Bowser L., Carninci P., Chang E., Dale J.M., Goldsmith A.D.,
 RA Hayashizaki Y., Ishida J., Jones T., Kamiya A., Karlin-Neumann G.,
 RA Kawai J., Lam B., Lee J.M., Lin J., Miranda M., Narusaka M.,
 RA Nguyen M., Onodera C.S., Palm C.J., Quach H.L., Sakurai T., Satou M.,
 RA Seki M., Southwick A., Tang C.C., Toriumi M., Wu H.C., Yamada K.,
 RA Ecker J.R.

RA Yamamura Y., Yu G., Yu S., Shinozaki K., Davis R.W., Theologis A.,
 RA Ecker J.R.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 RN [8]
 RP SEQUENCE FROM N.A.
 RA PubMed=14731261; DOI=10.1046/j.1365-313X.2003.01971.x;
 RA Meister R.J., Williams L.A., Monfared M.M., Gallagher T.L.,
 RA Kraft E.A., Nelson C.G., Gasser C.S.;
 RT "Definition and interactions of a positive regulatory element of the
 RT Arabidopsis INNER NO OUTER promoter.";
 RL Plant J. 37:426-438(2004).
 DR EMBL; AC006917; AAF79219.1; -
 DR EMBL; AC010657; AAF63172.1; -
 DR EMBL; AY058073; AAL24181.1; -
 DR EMBL; AY090305; AAL90966.1; -
 DR EMBL; AY380568; AAR25822.1; -
 DR InterPro; IPR010409; DUF1004.
 DR Pfam; PF06217; DUF1004; 1.
 SQ SEQUENCE 279 AA; 31168 MW; BAB996037E04D372 CRC64;
 Query Match 68.9%; Score 45.5; DB 2; Length 279;
 Best Local Similarity 77.8%; Pred. No. 50;
 Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1 CY-WGCGYW 8
 Db 200 CYRWGCGG 208
 ID Q6UIZ7 PRELIMINARY; PRT; 283 AA.
 AC Q6UIZ7;
 DT 05-JUL-2004 (TrEMBLrel. 27, Created)
 DT 05-JUL-2004 (TrEMBLrel. 27, Last sequence update)
 DE Basic pentacycysteine 1.
 GN Name=BPC1;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA PubMed=14731261; DOI=10.1046/j.1365-313X.2003.01971.x;
 RA Meister R.J., Williams L.A., Monfared M.M., Gallagher T.L.,
 RA Kraft E.A., Nelson C.G., Gasser C.S.;
 RT "Definition and interactions of a positive regulatory element of the
 RT Arabidopsis INNER NO OUTER promoter.";
 RL Plant J. 37:426-438(2004).
 DR EMBL; AY380334; AAR28441.1; -
 DR InterPro; IPR010409; DUF1004.
 DR Pfam; PF06217; DUF1004; 1.
 SQ SEQUENCE 283 AA; 31635 MW; BE69E04695CF20FC CRC64;
 Query Match 68.9%; Score 45.5; DB 2; Length 283;
 Best Local Similarity 77.8%; Pred. No. 50;
 Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1 CY-WGCGYW 8
 Db 204 CYRWGCGG 212
 ID Q9SKD0 PRELIMINARY; PRT; 283 AA.
 AC Q9SKD0;
 DT 01-MAY-2000 (TrEMBLrel. 13, Created)
 DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
 DE 05-JUL-2004 (TrEMBLrel. 27, Last annotation update)
 DE Hypothetical protein At2g01930 (Expressed protein).

GN Name=At2g01930;
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
 RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
 RA Carrincci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,
 RA Seki M., Shinn P., Southwick A., Shinzaki K., Davis R.W., Ecker J.R.,
 RA Theologis A.;
 RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Lin X., Kaul S., Shea T.P., Fujii C.Y., Shen M., VanAken S.E.,
 RA Barnstead M.E., Mason T.M., Bowman C.L., Ronning C.M., Benito M.-I.,
 RA Carrera A.J., Creasy T.H., Buell C.R., Town C.D., Nierman W.C.,
 RA Fraser C.M., Venter J.C.;
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP SEQUENCE FROM N.A.
 RA Town C.D., Kaul S.;
 RL Submitted (FEB-2002) to the EMBL/GenBank/DBJ databases.
 RN [4]
 RP SEQUENCE FROM N.A.
 RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
 RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
 RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
 RA Carrincci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
 RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
 RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinzaki K.,
 RA Davis R.W., Ecker J.R., Theologis A.;
 RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY065225; AAL38701.1; -
 DR EMBL; AC007265; AAM15473.1; -
 DR EMBL; AY096550; AAM20200.1; -
 DR PIR; H84430; H84430.
 DR InterPro; IPR010409; DUF1004.
 DR Pfam; PF06217; DUF1004; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 283 AA; 31648 MW; BAFBFCBD92362E07 CRC64;
 Query Match 68.9%; Score 45.5; DB 2; Length 283;
 Best Local Similarity 77.8%; Pred. No. 50;
 Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1 CY-WGCGYW 8
 Db 204 CYRWGCGW 212
 ID QBL9F3 PRELIMINARY; PRT; 285 AA.
 AC QBL9F3
 DT 01-OCT-2002 (TrEMBLrel. 22, Created)
 DT 01-OCT-2002 (TrEMBLrel. 22, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Hypothetical protein.
 OS Arabidopsis thaliana (Mouse-ear cress).
 OC Eukaryota; Viridiplantae; Streptophyta; Tracheophyta;
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
 OX NCBI_TaxID=3702;
 RN [1]
 RP SEQUENCE FROM N.A.
 RX Haas B.J., Volfovsky N., Town C.D., Troukhan M., Alexandrov N.,

RA Feldmann K.A., Flavell R.B., White O., Salzberg S.L.;
 RT "Full-length messenger RNA sequences greatly improve genome
 annotation.";
 RL Genome Biol. 3:RESEARCH0029-RESEARCH0029(2002).
 RN [2]
 RP SEQUENCE FROM N.A.
 RA Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,
 RA Feldmann K.;
 RL Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AY088464; AAM66000.1; -
 DR InterPro; IPR010409; DUF1004.
 DR Pfam; PF06217; DUF1004; 1.
 KW Hypothetical protein.
 SQ SEQUENCE 285 AA; 31887 MW; B97DB307A81CC504 CRC64;
 Query Match 68.9%; Score 45.5; DB 2; Length 285;
 Best Local Similarity 77.8%; Pred. No. 51;
 Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
 QY 1 CY-WGCGYW 8
 Db 206 CYRWGCGW 214
 ID Q7MYM7 PRELIMINARY; PRT; 463 AA.
 AC Q7MYM7
 DT 01-MAR-2004 (TrEMBLrel. 26, Created)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last sequence update)
 DT 01-MAR-2004 (TrEMBLrel. 26, Last annotation update)
 DE Probable transport protein YifK.
 GN NamesyifK; OrderedLocustNames=plu4650;
 OS Photorhabdus luminescens (subsp. laumondii).
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;
 OC Enterobacteriaceae; Photorhabdus.
 OX NCBI_TaxID=141679;
 RN [1]
 RP SEQUENCE FROM N.A.
 RC STRAIN=TT01;
 RX MEDLINE=22957627; PubMed=14528314;
 RA Duchaud E., Rusniok C., Frangeul L., Buchrieser C., Givaudan A.,
 RA Taourit S., Bocs S., Boursaux-Eude C., Chandler M., Charles J.-F.,
 RA Dassa E., Derose R., Derzelle S., Freyssinet G., Gaudriault S.,
 RA Medigue C., Lanois A., Powell K., Sigvier P., Vincent R., Wingate V.,
 RA Zouine M., Glaser P., Boemare N., Danchin A., Kunst F.;
 RT "The genome sequence of the entomopathogenic bacterium Photorhabdus
 luminescens.";
 RL Nat. Biotechnol. 21:1307-1313(2003).
 DR EMBL; BX571874; CAE17022.1; -
 DR PhotoList; plu4650;
 DR GO; GO:0016021; C:integral to membrane; IEA.
 DR GO; GO:0005279; F:amino acid-polyamine transporter activity; IEA.
 DR GO; GO:0006865; P:amino acid transport; IEA.
 DR GO; GO:0006810; P:transport; IEA.
 DR InterPro; IPR002293; AA/rel_permease1.
 DR InterPro; IPR004840; AAC_permease.
 DR InterPro; IPR004841; Permease_region.
 DR Pfam; PF00324; AA_permease; 1.
 DR PROSITE; PS00218; AMINO ACID PERMEASE 1; 1.
 KW Complete proteome; Transmembrane; Transport.
 SQ SEQUENCE 463 AA; 50838 MW; 317CD0ADCF4FCF25 CRC64;
 Query Match 68.2%; Score 45; DB 2; Length 463;
 Best Local Similarity 63.6%; Pred. No. 90;
 Matches 7; Conservative 0; Mismatches 0; Indels 4; Gaps 1;
 QY 2 YWGC---GYW 8
 Db 92 YWGLUTAWGYW 102

Search completed: April 18, 2005, 19:58:36

Job time : 73.2558 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 19:18:17 ; Search time 82.0465 Seconds
(without alignments)
37.711 Million cell updates/sec

Title: SEQ5

Perfect score: 66

Sequence: 1 cywgcgyw 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 2105692 seqs, 386760381 residues

Total number of hits satisfying chosen parameters: 2105692

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : A_Geneseq_16Dec04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001s:*
- 5: Geneseq2002s:*
- 6: Geneseq2003as:*
- 7: Geneseq2003bs:*
- 8: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	66	100.0	8	2	AAY03715 Fluorine-
2	66	100.0	8	3	AAY76817 Immunogen
3	66	100.0	8	7	ADG94005 Immunogen
4	66	100.0	8	8	ADL98014 Peptide h
5	48	72.7	54	3	AAY66023 Teioneras
6	48	72.7	249	7	ADC33110 Human nov
7	47	71.2	841	3	AAI18511 H2 homolo
8	46	69.7	14	8	ADQ11552 Myostatin
9	46	69.7	50	8	ADQ11655 Myostatin
10	46	69.7	265	6	ABU33201 Protein e
11	45.5	68.9	233	3	AAG40353 Arabidops
12	45.5	68.9	238	3	AAG60115 Arabidops
13	45.5	68.9	252	3	AAG05644 Arabidops
14	45.5	68.9	252	3	AAG40352 Arabidops
15	45.5	68.9	252	3	AAG38268 Arabidops
16	45.5	68.9	257	3	AAG60114 Arabidops
17	45.5	68.9	279	3	AAG38267 Arabidops
18	45.5	68.9	279	3	AAG05643 Arabidops
19	45.5	68.9	279	3	AAG40351 Arabidops
20	45.5	68.9	282	3	AAG38266 Arabidops
21	45.5	68.9	285	3	AAG60113 Arabidops
22	45.5	68.9	286	3	AAG05642 Arabidops
23	45	68.2	8	2	AAY03714 Fluorine-
24	45	68.2	8	2	AAY03716 Fluorine-
25	45	68.2	8	7	ADG94002 Iodinated

26	45	68.2	8	8	ADL98013 Peptide h
27	45	68.2	9	3	AAY76816 Immunogen
28	45	68.2	60	5	ABB57406 Human sec
29	45	68.2	75	4	AAO10292 Human pol
30	45	68.2	444	8	ADS42831 Bacterial
31	45	68.2	445	8	ADN17810 Bacterial
32	45	68.2	464	6	ABM68023 Photornab
33	43	65.2	60	4	ABB16263 Human ner
34	42	63.6	7	2	AAY42908 Somatosta
35	42	63.6	20	5	AAU90545 Insulin/i
36	42	63.6	115	3	AAB42716 Human ORF
37	42	63.6	332	2	AAW38424 Human cel
38	42	63.6	332	2	AAW38425 Rat cell
39	42	63.6	332	2	AAY01784 A human c
40	42	63.6	332	7	ADG62871 Human Pro
41	42	63.6	332	7	ADG62869 Rat Prote
42	42	63.6	332	7	ADG62867 Human Pro
43	42	63.6	332	7	ADG62865 Rat Prote
44	42	63.6	332	8	ADP12617 Protein e
45	42	63.6	345	4	AAG75458 Human col

ALIGNMENTS

RESULT 1

AAAY03715
ID AAY03715 standard; peptide; 8 AA.

AC AAY03715;

XX 08-JUN-1999 (first entry)

XX Fluorine-18 (F-18) labeled peptide 2.

XX 18F radionuclide; targeting vector; positron emission tomography; F-18;
KW radiolabeling; thiol; fluorine-18.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "N-terminal acetylation; optionally has a free or
protected thiol group"

FT Misc-difference 2 /note= "D-form residue"

FT Misc-difference 3 /note= "D-form residue"

FT Misc-difference 5 /note= "D-form residue"

FT Misc-difference 7 /note= "D-form residue; optionally has a free or
protected thiol group"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

XX The invention relates to a method for incorporating 18F radionuclide into
 CC peptide-containing targeting vectors for use in clinical positron
 CC emission tomography. Radiolabeling thiol-containing peptides with
 CC fluorine-18 (F-18) comprises reacting a peptide comprising a free thiol
 CC group with a labeling reagent of formula: 18F-(CH₂)_n-CR1R2-(CH₂)_n-X, or a
 CC fluorinated alkene in which at least one of the two double bonded carbon
 CC atoms bears at least one leaving group comprising I, Br, Cl, azide,
 CC tosylate, mesylate, nosylate or triflate. n, m = 0-2; n+m = 0-2; X = I,
 CC Br, Cl, azide, tosylate, mesylate, nosylate, triflate, maleimide
 CC (optionally substituted by 1-2 alkyl) or 3-sulfonamide; R1, R2 = I,
 CC Br, Cl, azide, tosylate, mesylate, nosylate, triflate, H, CONH₂, COOH,
 CC OH, sulfonic acid, tertiary amine, quaternary ammonium, alkyl (optionally
 CC substituted by CONH₂, COOH, OH, sulfonic acid, tertiary amine or
 CC quaternary ammonium), COR', CONR'2 or COR'; and R' = 1-6C alkyl or
 CC phenyl. The method is used for radiolabeling peptide-containing targeting
 CC vectors such as proteins, antibodies, antibody fragments and receptor-
 CC targeted peptides for use in routine clinical positron emission
 CC tomography. The method is simple and efficient. The method uses the
 CC unique property of the free thiol groups which are rapidly alkylated at
 CC neutral pH and moderate temperature. Sequences AAY03714-716 represent
 CC examples of F-18 labeled peptides used in the method of detecting a
 CC tissue
 CC
 CC Sequence 8 AA;
 CC

Query Match 100.0%; Score 66; DB 2; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CYWGGCGYW 8
 |||||
 DB 1 CYWGGCGYW 8

RESULT 2

AAY76817
 ID AAY76817 standard; peptide; 8 AA.

XX AC AAY76817;
 XX

DT 28-APR-2000 (first entry)

XX Immunogenic peptide for bi-specific antibody recognition.

XX Immunogenic peptide; bi-specific antibody; diagnosis; immune response;
 KW diseased tissue identification; therapy.

XX Synthetic.

XX Key Location/Qualifiers

FT Misc-difference 1 /note= "acetylated; modified with free amino acid group,
 FT protected amino acid group, chelating agent or a metal-
 FT chelate complex"

FT Misc-difference 2 /note= "D-form residue"

FT Misc-difference 3 /note= "D-form residue"

FT Misc-difference 5 /note= "D-form residue"

FT Misc-difference 7 /note= "D-form residue; modified with free amino acid
 FT group, protected amino acid group, chelating agent or a
 FT metal-chelate complex"

FT Misc-difference 7 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

XX WO966951-A2.

XX 29-DEC-1999.

XX 22-JUN-1999; 99WO-US013879.

PF

XX PR
 PR 22-JUN-1998; 98US-0090142P.
 PR 14-OCT-1998; 98US-0104158P.

XX (IMMU-) IMMUNOMEDICS INC.
 XX Hansen HJ, Griffiths GL, Leung S, McBride WJ, Qu Z;

XX WPI; 2000-160561/14.

XX Bi-specific antibodies that bind specific target tissue and targeted
 CC conjugates.
 XX Claim 22; Page 61; 76pp; English.

XX This sequence represents an immunogenic peptide for a bi-specific
 CC antibody. The invention relates to a method of treating or identifying
 CC diseased tissues in a patient comprising administering a bi-specific
 CC antibody (or fragment) having at least 1 arm (A) that specifically
 CC binds a targeted tissue and at least 1 arm (B) that specifically binds a
 CC targetable conjugate. The methods and bi-specific antibodies and fusion
 CC proteins are useful for pre-targeting methods of diagnosis and therapy.
 CC It is advantageous to raise bi-specific antibodies against a targetable
 CC conjugate that is capable of carrying at least 1 diagnostic or
 CC therapeutic agent. The characteristics of the chelator, metal chelate
 CC complex, therapeutic agent or diagnostic agent can be varied to
 CC accommodate differing applications without raising new bi-specific
 CC antibodies for each new application. The targetable conjugate is selected
 CC to elicit sufficient immune responses and also for rapid in vivo
 CC clearance when used within the bi-specific antibody targeting method
 CC

XX Sequence 8 AA;

Query Match 100.0%; Score 66; DB 3; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CYWGGCGYW 8
 |||||
 DB 1 CYWGGCGYW 8

RESULT 3

ADG94005
 ID ADG94005 standard; peptide; 8 AA.

XX AC ADG94005;
 XX

XX 11-MAR-2004 (first entry)

XX Immunogenic peptide.

XX Immunogenic peptide; multi-specific antibody; polymer conjugate; tumour;
 KW cytostatic; photodynamic therapy.

XX Synthetic.

XX Key Location/Qualifiers

FT Modified-site 1 /note= "Optionally methylated or Acetylated"

FT Misc-difference 2 /note= "D-form residue"

FT Misc-difference 3 /note= "D-form residue"

FT Misc-difference 5 /note= "D-form residue"

FT Misc-difference 5 /note= "optionally methylated or Acetylated D-form
 FT residue"

FT Misc-difference 7 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

FT Misc-difference 8 /note= "D-form residue"

XX US2003026764-A1.

XX PD 06-FEB-2003.
 XX PF 31-JUL-2002; 2002US-00209592.
 XX PR 31-JUL-2001; 2001US-0308605P.
 XX PA (IMMU-) IMMUNOMEDICS INC.
 XX PI Griffiths GL;
 XX DR WPI; 2003-801085/75.
 XX PT Targeting an agent towards a target site in a tissue, by administering in
 XX PT a tissue, a multi-specific antibody or its fragment and a polymer
 XX PT conjugate that binds to the capture arm of the multi-specific antibody.
 XX PS Disclosure; Page 14; 19pp; English.
 XX CC The invention relates to targeting an agent towards a target site in a
 CC tissue, comprising administering to the tissue, a multi-specific antibody
 CC or its fragment, comprising a targeting arm that binds to an antigen of
 CC the target site and a capture arm that binds to a polymer conjugate, and
 CC administering a polymer conjugate that binds to the capture arm, the
 CC conjugate has a polymer conjugated to the agent such as therapeutic
 CC agent, a peptide, an enzyme and a labelled ligand. Also included is a kit
 CC useful for targeting a target site within a tissue in a subject or tissue
 CC sample comprising the above mentioned multi-specific antibody or its
 CC fragment and a polymer conjugate. The method is used for targeting an
 CC agent towards a target site in a tissue (e.g. a tumour). The method is
 CC also useful for therapeutic or diagnostic purposes and further in
 CC photodynamic therapy. The present sequence is an immunogenic peptide used
 CC in the method of the invention.
 XX SQ Sequence 8 AA;
 Query Match 100.0%; Score 66; DB 7; Length 8;
 Best Local Similarity 100.0%; Pred. No. 1.8e+06;
 Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 CYWGCGYW 8
 Db 1 CYWGCGYW 8
 RESULT 4
 ADL98014
 ID ADL98014 standard; peptide; 8 AA.
 XX AC ADL98014;
 XX DT 20-MAY-2004 (first entry)
 XX DE Peptide hapten #2.
 XX KW photodynamic diagnosis; cancer; tumour; cardiovascular lesion;
 KW inflammatory disease; neurodegenerative disease; metabolic disease;
 KW infectious disease; B-cell malignancy; Alzheimer's disease; amyloidosis;
 KW autoimmune disease; bacterial infection; fungal infection;
 KW parasitic infection; viral infection;
 KW carcinoembryonic antigen-expressing tumour.
 XX OS Synthetic.
 XX PH Key Location/Qualifiers
 FT Modified-site 1
 FT /note= "Optionally Cysteinylyl; if Cys the residue is
 FT methylated; N-terminal acetylated."
 FT Misc-difference 2
 FT /note= "D form residue"
 FT Misc-difference 3
 FT /note= "D form residue"
 FT Modified-site 5

FT FT /note= "Optionally Cysteinylyl; if Cys the residue is
 FT methylated; N-terminal acetylated."
 FT Misc-difference 7
 FT /note= "D form residue"
 FT Misc-difference 8
 FT /note= "D form residue"
 XX XX US2004043030-A1.
 XX PD 04-MAR-2004.
 XX PF 09-JUN-2003; 2003US-00456580.
 XX PR 31-JUL-2001; 2001US-0308605P.
 XX PR 31-JUL-2002; 2002US-00209592.
 XX PA (IMMU-) IMMUNOMEDICS INC.
 XX PI Griffiths GL, Goldenberg DM, Hansen HJ;
 XX WPI; 2004-313738/29.
 XX PT Treating cancer and metabolic diseases by administering a multi-specific
 XX PT antibody having a targeting arm that binds to an antigen and a capture
 XX PT arm that binds to a polymer conjugate comprising a therapeutic agent.
 XX PS Disclosure; Page 16; 24pp; English.
 XX CC The invention relates to a method of diagnosing or treating a disease or
 CC disorder. The method involves administering to a tissue a multi-specific
 CC antibody (1) or antibody fragment, comprising a targeting arm that binds
 CC to an antigen on the target site, and a capture arm that binds to a
 CC polymer conjugate, and administering to the tissue a polymer conjugate
 CC that binds to the capture arm, the polymer conjugate comprising a polymer
 CC conjugated to a diagnostic or therapeutic agent. Also included is a
 CC method for photodynamic diagnosis or treatment of a disease or disorder;
 CC or intravascular or endoscopic method for diagnosing or treating a
 CC disease or disorder. The method is useful for diagnosing or treating a
 CC disease or disorder chosen from cancer (oesophageal, gastric, colonic,
 CC rectal, pancreatic, lung, breast, ovarian, urinary bladder, endometrial,
 CC cervical, testicular, renal, adrenal and liver cancer, solid tumour, B-
 CC cell malignancy or T-cell malignancy); cardiovascular lesion; an
 CC inflammatory disease; neurodegenerative disease; metabolic disease; and
 CC an infectious disease. The B-cell malignancy is chosen from indolent
 CC forms of B-cell lymphomas, aggressive forms of B-cell lymphomas, chronic
 CC lymphatic leukaemias, acute lymphatic leukaemias, and multiple myeloma.
 CC The solid tumour is chosen melanoma, carcinoma (preferably renal
 CC carcinoma, lung carcinoma, intestinal carcinoma, and stomach carcinoma),
 CC glioma and sarcoma. The cardiovascular lesion is chosen from infarct,
 CC clot, embolus, atherosclerotic plaque and ischaemia. The
 CC neurodegenerative disease is Alzheimer's disease. The metabolic disease
 CC is amyloidosis, where the antibody binds amyloid. The disease or disorder
 CC is displaced or ectopic normal tissue chosen from endometrium, thymus,
 CC spleen and parathyroid. The method can be used for normal tissue
 CC ablation, where the tissue is chosen from bone marrow and spleen. The
 CC disease or disorder is an autoimmune disease such as myasthenia gravis,
 CC lupus nephritis, lupus erythematosus, and rheumatoid arthritis, Class III
 CC autoimmune diseases such as immune-mediated thrombocytopenias, such as
 CC acute idiopathic thrombocytopenic purpura and chronic idiopathic
 CC thrombocytopenic purpura, dermatomyositis, Sjogren's syndrome, multiple
 CC sclerosis, Sydenham's chorea, myasthenia gravis, systemic lupus
 CC erythematosus, lupus nephritis, rheumatic fever, polyglandular syndromes,
 CC bullous pemphigoid, diabetes mellitus, Henoch-Schönlein purpura, post-
 CC streptococcal nephritis, erythema nodosum, Takayasu's arteritis,
 CC Addison's disease, rheumatoid arthritis, sarcoidosis, ulcerative colitis,
 CC erythema multiforme, IGA nephropathy, polyarteritis nodosa, ankylosing
 CC spondylitis, Goodpasture's syndrome, thromboangiitis obliterans, primary
 CC biliary cirrhosis, Hashimoto's thyroiditis, thyrotoxicosis, scleroderma,
 CC chronic active hepatitis, polymyositis/dermatomyositis, polychondritis,
 CC pemphigus vulgaris, Wegener's granulomatosis, membranous nephropathy,
 CC amyotrophic lateral sclerosis, tabes dorsalis, giant cell
 CC arteritis/polyneuritis, pernicious anaemia, rapidly progressive
 CC glomerulonephritis, or fibrosing alveolitis. The infectious disease is

chosen from bacterial, fungal, parasitic and viral lesion. The infectious disease is caused by a fungus chosen from Microsporium, Trichophyton, Epidermophyton, Sporothrix schenckii, Cryptococcus neoformans, Coccidioides immitis, Histoplasma capsulatum, Blastomyces dermatitidis, and Candida albicans. The infectious disease is caused by a virus chosen from HIV, herpes virus, cytomegalovirus, rabies virus, influenza virus, hepatitis B virus, Sendai virus, feline leukemia virus, Reo virus, polio virus, human serum parvo-like virus, simian virus 40, respiratory syncytial virus, mouse mammary tumour virus, Varicella-Zoster virus, Dengue virus, rubella virus, measles virus, adenovirus, human T-cell leukemia viruses, Epstein-Barr virus, murine leukemia virus, mumps virus, vesicular stomatitis virus, Sindbis virus, lymphocytic choriomeningitis virus, wart virus and blue tongue virus. The infectious disease is caused by a bacterium chosen from Bacillus anthracis, Streptococcus agalactiae, Legionella pneumophila, Streptococcus pyogenes, Escherichia coli, Neisseria gonorrhoeae, Neisseria meningitidis, Pneumococcus, Haemophilus influenzae B, Treponema pallidum, Lyme disease spirochetes, Pseudomonas aeruginosa, Mycobacterium leprae, Brucella abortus, Mycobacterium tuberculosis, and Tetanus toxin. The infectious disease is caused by a protozoa chosen from Plasmodium falciparum, Plasmodium vivax, Toxoplasma gondii, Trypanosoma rangeli, Trypanosoma cruzi, Trypanosoma rhodesiense, Trypanosoma brucei, Schistosoma mansoni, Schistosoma japonicum, Babesia bovis, Eimeria tenella, Onchocerca volvulus, Leishmania tropica, Trichinella spiralis, Onchocerca volvulus, Theileria parva, Taenia hydatigena, Taenia ovis, Taenia saginata, Echinococcus granulosus, and Mesocystoides corti. The infectious disease is caused by a mycoplasma chosen from Mycoplasma arthritidis, M. hyorhinis, M. orale, M. arginini, Acholeplasma laidlawii, M. salivarium and M. pneumoniae. The cancer is preferably chosen from carcinoembryonic antigen (CEA) - expressing tumour or a CD20-expressing malignancy. The present sequence represents a peptide used in the method of the invention.

XX Sequence 8 AA;

Query Match 100.0%; Score 66; DB 8; Length 8;
Best Local Similarity 100.0%; Pred. No. 1.8e+06;
Matches 8; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CYWGGCYW 8
Db 1 CYWGGCYW 8

RESULT 5
AAY66023
ID AAY66023 standard; peptide; 54 AA.

XX AAY66023;
AC
AC 10-FEB-2000 (first entry)
DT Telomerase associated protein TP-1 mutant peptide 6.
DE
XX Human; frameshift mutant; T cell response; tumour; treatment; cancer;
KW mutin.
KW
XX Homo sapiens.
OS Synthetic.
OS
XX WO958552-A2.
PN
XX 18-NOV-1999.
PD
XX 03-MAY-1999; 99WO-NO000143.
PF
XX 08-MAY-1998; 98NO-00002097.
PR
XX (NHSD) NORSE HYDRO AS.
PA
XX Gaudernack G, Eriksen JA, Moller M, Gjertsen MK, Saeterdal I;
PI WPI; 2000-039064/03.
XX

PT New peptides derived from genes with frameshift mutations, used to develop products for the treatment and prophylaxis of cancers.
PT
XX Claim 13; Page 36; 166pp; English.

XX Peptides AAY65684-Y66142 are fragments of mutant proteins arising from a frameshift mutation in a gene from a cancer cell. The peptides are characterised in that they: (i) are at least 8 amino acids long and a fragment of a mutant protein arising from a frameshift mutation in a gene of a cancer cell; (ii) consist of at least one amino acid of the mutant part of a protein sequence encoded by the gene; (iii) comprise 0-10 amino acid from the carboxyl terminus of the normal part of the protein sequence preceding the amino terminus of the mutant sequence and may further extend to the carboxyl terminus of the mutant sequence and as determined by a new stop codon generated by the frameshift mutation; and (iv) induce, either in their full lengths or after processing by an antigen presenting cell (APC), T cell responses. The genes that the peptides are derived from, are characterised as susceptible to frameshift mutation by having a mono nucleoside base repeat sequence of at least 5 residues, or a di-nucleoside base repeat sequence of at least 4 di-nucleoside base units. The peptides are created by the addition or deletion of 1 or 2 nucleoside base residues from the repeat sequence. The novel peptides can elicit T cell responses and toxicity against tumours and cancer cells carrying genes with frameshift mutations. The novel peptides and DNA sequences can be used for the preparation of a composition for the treatment or prophylaxis of cancer

XX Sequence 54 AA;

Query Match 72.7%; Score 48; DB 3; Length 54;
Best Local Similarity 75.0%; Pred. No. 32;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CYWGGCYW 8
Db 5 CEWGGCWS 12

RESULT 6
ADC33110
ID ADC33110 standard; protein; 249 AA.
XX
XX ADC33110;
AC
AC 18-DEC-2003 (first entry)
DT
DE Human novel contig-encoded polypeptide sequence, SEQ ID NO:3192.
XX

KW Human; diagnostic; drug screening; forensics; gene mapping;
KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
KW ulcers; osteoporosis; autoimmune disease; cancer;
KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
KW neuroprotective; anti-nausea; anticoagulant; thrombolytic; vulvar;
KW antitumor; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
KW gene therapy.

XX Homo sapiens.
OS
XX WO2003029271-A2.
PN
XX 10-APR-2003.
PD
XX 24-SEP-2002; 2002WO-US030474.
PF
XX 24-SEP-2001; 2001US-0324631P.
PR
XX (HYSE-) HYSEQ INC.

XX Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
PI Haley-Vicente D, Drmanac RT;
XX

DR WPI; 2004-525896/50.

XX New binding agent which inhibits myostatin, useful for treating muscular

PT dystrophy, amyotrophic lateral sclerosis, congestive obstructive

PT pulmonary disease, stroke, aging, diabetes, obesity, osteoporosis.

XX Example 1; SEQ ID NO 33; 287pp; English.

XX

XX The present invention relates to novel binding agents comprising at least

CC one peptide capable of binding myostatin and inhibiting its activity. In

CC one embodiment, the binding agent comprises at least one myostatin-

CC binding peptide attached directly or indirectly to at least one vehicle

CC such as a polymer or an Fc domain. Myostatin (also known as

CC growth/differentiation factor 8, GDF-8) is a transforming growth factor-

CC beta (TGF-beta) family member known to be involved in regulation of

CC skeletal muscle mass. The binding agents increase lean muscle mass when

CC administered to animals and decrease fat to muscle ratios. The binding

CC agents are useful for treating muscle-wasting disease, e.g. muscular

CC dystrophy, amyotrophic lateral sclerosis, congestive obstructive

CC pulmonary disease, chronic heart failure, cancer, AIDS, renal failure,

CC uremia, rheumatoid arthritis, age-related sarcopenia, muscle-wasting due

CC to prolonged bedrest, spinal chord injury, stroke, bone fracture, or

CC aging, and myosin-related metabolic disorder, e.g. diabetes, obesity,

CC hyperglycaemia, bone loss, or osteoporosis. The present sequence is one

CC such myostatin-binding peptide.

XX

XX Sequence 14 AA;

QY Query Match 69.7%; Score 46; DB 8; Length 14;

DB Best Local Similarity 71.4%; Pred. No. 19;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 YMGCGYW 8

DB : || || ||

1 HWACGYW 7

RESULT 9

ADQ11655

ID ADQ11655 standard; peptide; 50 AA.

AC ADQ11655;

XX

XX 07-OCT-2004 (first entry)

XX

XX Myostatin binding peptide, Myostatin-2xTNF-23 kc, SEQ ID 136.

XX

XX Muscular; Muscular; Gene Therapy; Antidiabetic; Anorectic; Myostatin;

XX myostatin-binding peptide; growth/differentiation factor 8; GDF-8;

XX skeletal muscle mass; muscle-wasting disease; muscular dystrophy;

XX amyotrophic lateral sclerosis; congestive obstructive pulmonary disease;

XX chronic heart failure; cancer; AIDS; renal failure; uremia;

XX rheumatoid arthritis; age-related sarcopenia; muscle-wasting;

XX spinal chord injury; stroke; bone fracture; aging; diabetes; obesity;

XX hyperglycaemia; bone loss; osteoporosis.

XX Unidentified.

OS

XX WO2004058988-A2.

XX

XX 15-JUL-2004.

XX

XX 19-DEC-2003; 2003WO-US040781.

XX

XX 20-DEC-2002; 2002US-0435923P.

XX

XX (AMGE-) AMGEN INC.

XX

XX Han H, Min H, Boone TC;

XX

XX WPI; 2004-525896/50.

XX

XX New binding agent which inhibits myostatin, useful for treating muscular

PT dystrophy, amyotrophic lateral sclerosis, congestive obstructive

PT pulmonary disease, stroke, aging, diabetes, obesity, osteoporosis.

PS Example 1; SEQ ID NO 136; 287pp; English.

XX

XX The present invention relates to novel binding agents comprising at least

CC one peptide capable of binding myostatin and inhibiting its activity. In

CC one embodiment, the binding agent comprises at least one myostatin-

CC binding peptide attached directly or indirectly to at least one vehicle

CC such as a polymer or an Fc domain. Myostatin (also known as

CC growth/differentiation factor 8, GDF-8) is a transforming growth factor-

CC beta (TGF-beta) family member known to be involved in regulation of

CC skeletal muscle mass. The binding agents increase lean muscle mass when

CC administered to animals and decrease fat to muscle ratios. The binding

CC agents are useful for treating muscle-wasting disease, e.g. muscular

CC dystrophy, amyotrophic lateral sclerosis, congestive obstructive

CC pulmonary disease, chronic heart failure, cancer, AIDS, renal failure,

CC uremia, rheumatoid arthritis, age-related sarcopenia, muscle-wasting due

CC to prolonged bedrest, spinal chord injury, stroke, bone fracture, or

CC aging, and myosin-related metabolic disorder, e.g. diabetes, obesity,

CC hyperglycaemia, bone loss, or osteoporosis. The present sequence is one

CC such myostatin-binding peptide.

XX

XX Sequence 50 AA;

QY Query Match 69.7%; Score 46; DB 8; Length 50;

DB Best Local Similarity 71.4%; Pred. No. 55;

Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 YMGCGYW 8

DB : || || ||

1 HWACGYW 7

RESULT 10

ABU33201

ID ABU33201 standard; protein; 265 AA.

XX

XX AC ABU33201;

XX

XX 19-JUN-2003 (first entry)

XX

XX Protein encoded by Prokaryotic essential gene #18728.

XX

XX Antisense; prokaryotic essential gene; cell proliferation; drug design.

XX

XX Legionella pneumophila.

OS

XX WO200277183-A2.

XX

XX 03-OCT-2002.

PD

XX

XX 21-MAR-2002; 2002WO-US009107.

PF

XX

XX 21-MAR-2001; 2001US-00815242.

PR

XX 06-SEP-2001; 2001US-00948993.

PR

XX 25-OCT-2001; 2001US-0342923P.

PR

XX 08-FEB-2002; 2002US-00072851.

PR

XX 06-MAR-2002; 2002US-0362699P.

PR

XX (ELIT-) ELITRA PHARM INC.

PA

XX

XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohleen KL, Zyskind JW;

PI Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;

PI

XX WPI; 2003-029926/02.

DR

XX N-PSDB; ACA37071.

XX

XX New antisense nucleic acids, useful for identifying proteins or screening

PT for homologous nucleic acids required for cellular proliferation to

PT isolate candidate molecules for rational drug discovery programs.

XX

XX Claim 25; SEQ ID NO 61125; 1766pp; English.

XX The invention relates to an isolated nucleic acid comprising any one of
CC the 6213 antisense sequences given in the specification where expression
CC of the nucleic acid inhibits proliferation of a cell. Also included are:
CC (1) a vector comprising a promoter operably linked to the nucleic acid
CC encoding a polypeptide whose expression is inhibited by the antisense
CC nucleic acid; (2) a host cell containing the vector; (3) an isolated
CC polypeptide or its fragment whose expression is inhibited by the
CC antisense nucleic acid; (4) an antibody capable of specifically binding
CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
CC proliferation or the activity of a gene in an operon required for
CC proliferation; (7) identifying a compound that influences the activity of
CC the gene product or that has an activity against a biological pathway
CC required for proliferation, or that inhibits cellular proliferation; (8)
CC identifying a gene required for cellular proliferation or the biological
CC pathway in which a proliferation-required gene or its gene product lies
CC or a gene on which the test compound that inhibits proliferation of an
CC organism acts; (9) manufacturing an antibiotic; (10) profiling a
CC compound's activity; (11) a culture comprising strains in which the gene
CC product is overexpressed or underexpressed; (12) determining the extent
CC to which each of the strains is present in a culture or collection of
CC strains; or (13) identifying the target of a compound that inhibits the
CC proliferation of an organism. The antisense nucleic acids are useful for
CC identifying proteins or screening for homologous nucleic acids required
CC for cellular proliferation to isolate candidate molecules for rational
CC drug discovery programs, or for screening homologous nucleic acids
CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,
CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is encoded by one of
CC the target prokaryotic essential genes. Note: The sequence data for this
CC patent did not form part of the printed specification, but was obtained
CC in electronic format directly from WIPO at
CC ftp.wipo.int/pub/published_pct_sequences

XX
SQ Sequence 265 AA;

Query Match 69.7%; Score 46; DB 6; Length 265;
Best Local Similarity 62.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CYWCGYW 8
|:|:|:
Db 65 CFWACGYF 72

RESULT 11
AAG40353
ID AAG40353 standard; protein; 233 AA.
XX AC AAG40353;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 50058.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX FN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.

06-APR-1999; 99US-0128234P.
08-APR-1999; 99US-0128714P.
16-APR-1999; 99US-0129845P.
19-APR-1999; 99US-0130077P.
21-APR-1999; 99US-0130449P.
23-APR-1999; 99US-0130510P.
28-APR-1999; 99US-0130891P.
30-APR-1999; 99US-0131449P.
30-APR-1999; 99US-0132048P.
04-MAY-1999; 99US-0132407P.
05-MAY-1999; 99US-0132484P.
06-MAY-1999; 99US-0132485P.
06-MAY-1999; 99US-0132486P.
07-MAY-1999; 99US-0132487P.
11-MAY-1999; 99US-0132863P.
11-MAY-1999; 99US-0134256P.
14-MAY-1999; 99US-0134218P.
14-MAY-1999; 99US-0134219P.
14-MAY-1999; 99US-0134221P.
14-MAY-1999; 99US-0134370P.
18-MAY-1999; 99US-0134768P.
19-MAY-1999; 99US-0134941P.
20-MAY-1999; 99US-0135124P.
21-MAY-1999; 99US-0135353P.
24-MAY-1999; 99US-0135629P.
25-MAY-1999; 99US-0136021P.
27-MAY-1999; 99US-0136392P.
28-MAY-1999; 99US-0136782P.
01-JUN-1999; 99US-0137222P.
03-JUN-1999; 99US-0137528P.
04-JUN-1999; 99US-0137502P.
07-JUN-1999; 99US-0137724P.
08-JUN-1999; 99US-0138094P.
10-JUN-1999; 99US-0138540P.
10-JUN-1999; 99US-0138847P.
14-JUN-1999; 99US-0139119P.
16-JUN-1999; 99US-0139452P.
17-JUN-1999; 99US-0139453P.
17-JUN-1999; 99US-0139492P.
18-JUN-1999; 99US-0139454P.
18-JUN-1999; 99US-0139455P.
18-JUN-1999; 99US-0139456P.
18-JUN-1999; 99US-0139457P.
18-JUN-1999; 99US-0139458P.
18-JUN-1999; 99US-0139459P.
18-JUN-1999; 99US-0139460P.
18-JUN-1999; 99US-0139461P.
18-JUN-1999; 99US-0139462P.
18-JUN-1999; 99US-0139463P.
18-JUN-1999; 99US-0139750P.
18-JUN-1999; 99US-0139763P.
21-JUN-1999; 99US-0139817P.
22-JUN-1999; 99US-0139899P.
23-JUN-1999; 99US-0140353P.
23-JUN-1999; 99US-0140354P.
24-JUN-1999; 99US-0140695P.
28-JUN-1999; 99US-0140823P.
28-JUN-1999; 99US-0140991P.
30-JUN-1999; 99US-0141287P.
01-JUL-1999; 99US-0141842P.
01-JUL-1999; 99US-0142154P.
02-JUL-1999; 99US-0142055P.
06-JUL-1999; 99US-0142390P.
08-JUL-1999; 99US-0142803P.
09-JUL-1999; 99US-0142920P.
12-JUL-1999; 99US-0142977P.
13-JUL-1999; 99US-0143542P.
14-JUL-1999; 99US-0143624P.
15-JUL-1999; 99US-0144005P.
16-JUL-1999; 99US-0144085P.
16-JUL-1999; 99US-0144086P.
19-JUL-1999; 99US-0144325P.
19-JUL-1999; 99US-0144331P.

PR 28-APR-1999;	99US-0131449P.	PR 21-JUL-1999;	99US-0144814P.
PR 30-APR-1999;	99US-0132048P.	PR 21-JUL-1999;	99US-0145086P.
PR 30-APR-1999;	99US-01322407P.	PR 21-JUL-1999;	99US-0145088P.
PR 04-MAY-1999;	99US-01322484P.	PR 22-JUL-1999;	99US-0145085P.
PR 05-MAY-1999;	99US-01322485P.	PR 22-JUL-1999;	99US-0145087P.
PR 06-MAY-1999;	99US-01322486P.	PR 22-JUL-1999;	99US-0145089P.
PR 06-MAY-1999;	99US-01322487P.	PR 22-JUL-1999;	99US-0145192P.
PR 07-MAY-1999;	99US-01322863P.	PR 23-JUL-1999;	99US-0145145P.
PR 11-MAY-1999;	99US-0134256P.	PR 23-JUL-1999;	99US-0145218P.
PR 14-MAY-1999;	99US-0134218P.	PR 23-JUL-1999;	99US-0145224P.
PR 14-MAY-1999;	99US-0134219P.	PR 26-JUL-1999;	99US-0145276P.
PR 14-MAY-1999;	99US-0134221P.	PR 27-JUL-1999;	99US-0145913P.
PR 14-MAY-1999;	99US-0134370P.	PR 27-JUL-1999;	99US-0145918P.
PR 18-MAY-1999;	99US-0134768P.	PR 27-JUL-1999;	99US-0145919P.
PR 19-MAY-1999;	99US-0134941P.	PR 28-JUL-1999;	99US-0145951P.
PR 20-MAY-1999;	99US-0135124P.	PR 02-AUG-1999;	99US-0146386P.
PR 21-MAY-1999;	99US-0135353P.	PR 02-AUG-1999;	99US-0146388P.
PR 24-MAY-1999;	99US-0135629P.	PR 02-AUG-1999;	99US-0146389P.
PR 25-MAY-1999;	99US-0136021P.	PR 03-AUG-1999;	99US-0147038P.
PR 27-MAY-1999;	99US-0136392P.	PR 04-AUG-1999;	99US-0147204P.
PR 28-MAY-1999;	99US-0136782P.	PR 04-AUG-1999;	99US-0147302P.
PR 01-JUN-1999;	99US-0137222P.	PR 05-AUG-1999;	99US-0147192P.
PR 03-JUN-1999;	99US-0137528P.	PR 05-AUG-1999;	99US-0147260P.
PR 04-JUN-1999;	99US-0137502P.	PR 06-AUG-1999;	99US-0147303P.
PR 07-JUN-1999;	99US-0137724P.	PR 06-AUG-1999;	99US-0147416P.
PR 08-JUN-1999;	99US-0138094P.	PR 09-AUG-1999;	99US-0147493P.
PR 10-JUN-1999;	99US-0138540P.	PR 09-AUG-1999;	99US-0147935P.
PR 10-JUN-1999;	99US-0138847P.	PR 10-AUG-1999;	99US-0148171P.
PR 14-JUN-1999;	99US-0139119P.	PR 11-AUG-1999;	99US-0148319P.
PR 16-JUN-1999;	99US-0139452P.	PR 12-AUG-1999;	99US-0148341P.
PR 16-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148565P.
PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 18-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149426P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 23-AUG-1999;	99US-0149930P.
PR 18-JUN-1999;	99US-0139462P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139463P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139763P.	PR 27-AUG-1999;	99US-0151066P.
PR 21-JUN-1999;	99US-0139817P.	PR 27-AUG-1999;	99US-0151080P.
PR 22-JUN-1999;	99US-0139899P.	PR 30-AUG-1999;	99US-0151303P.
PR 23-JUN-1999;	99US-0140353P.	PR 31-AUG-1999;	99US-0151438P.
PR 23-JUN-1999;	99US-0140354P.	PR 01-SEP-1999;	99US-0151930P.
PR 24-JUN-1999;	99US-0140695P.	PR 07-SEP-1999;	99US-0152363P.
PR 28-JUN-1999;	99US-0140823P.	PR 10-SEP-1999;	99US-0153070P.
PR 29-JUN-1999;	99US-0140991P.	PR 13-SEP-1999;	99US-0153758P.
PR 30-JUN-1999;	99US-0141287P.	PR 15-SEP-1999;	99US-0154018P.
PR 01-JUL-1999;	99US-0141842P.	PR 16-SEP-1999;	99US-0154039P.
PR 01-JUL-1999;	99US-0142154P.	PR 20-SEP-1999;	99US-0154779P.
PR 02-JUL-1999;	99US-0142055P.	PR 22-SEP-1999;	99US-0155139P.
PR 06-JUL-1999;	99US-0142390P.	PR 23-SEP-1999;	99US-0155486P.
PR 08-JUL-1999;	99US-0142803P.	PR 24-SEP-1999;	99US-0155659P.
PR 09-JUL-1999;	99US-0142920P.	PR 28-SEP-1999;	99US-0156458P.
PR 12-JUL-1999;	99US-0142877P.	PR 29-SEP-1999;	99US-0156596P.
PR 13-JUL-1999;	99US-0143842P.	PR 04-OCT-1999;	99US-0157117P.
PR 14-JUL-1999;	99US-0143624P.	PR 05-OCT-1999;	99US-0157753P.
PR 15-JUL-1999;	99US-0144005P.	PR 06-OCT-1999;	99US-0157865P.
PR 16-JUL-1999;	99US-0144085P.	PR 07-OCT-1999;	99US-0158029P.
PR 16-JUL-1999;	99US-0144086P.	PR 08-OCT-1999;	99US-0158232P.
PR 19-JUL-1999;	99US-0144325P.	PR 12-OCT-1999;	99US-0158369P.
PR 19-JUL-1999;	99US-0144331P.	PR 13-OCT-1999;	99US-0159293P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144333P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144334P.	PR 14-OCT-1999;	99US-0159329P.
PR 19-JUL-1999;	99US-0144335P.	PR 14-OCT-1999;	99US-0159330P.
PR 20-JUL-1999;	99US-0144352P.	PR 14-OCT-1999;	99US-0159331P.
PR 20-JUL-1999;	99US-0144632P.	PR 14-OCT-1999;	99US-0159637P.
PR 20-JUL-1999;	99US-0144844P.	PR 14-OCT-1999;	99US-0159638P.

```
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 23-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      68.9%; Score 45.5; DB 3; Length 238;
Best Local Similarity 77.8%; Pred. No. 2.3e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY      1 CY-WGCGYW 8
      |||||
Db      159 CYRWGCGGW 167

RESULT 13
AAG05644
ID AAG05644 standard; protein; 252 AA.
XX
AC AAG05644;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 2124.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
```

```
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145242P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145313P.
PR 27-JUL-1999; 99US-0145318P.
PR 27-JUL-1999; 99US-0145319P.
PR 28-JUL-1999; 99US-0145351P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.

PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 68.9%; Score 45.5; DB 3; Length 252;
Best Local Similarity 77.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CY-WGCGYW 8
Db 173 CYRWGCGW 181

RESULT 14
AAG40352
ID AAG40352 standard; protein; 252 AA.
XX AC AAG40352;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 50057.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0129845P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 23-APR-1999; 99US-0130891P.
XX PR 30-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 30-APR-1999; 99US-0132407P.
XX PR 04-MAY-1999; 99US-0132484P.
XX PR 05-MAY-1999; 99US-0132485P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 07-MAY-1999; 99US-0132487P.
XX PR 11-MAY-1999; 99US-0132863P.
XX PR 14-MAY-1999; 99US-0134256P.
XX PR 14-MAY-1999; 99US-0134218P.
XX PR 14-MAY-1999; 99US-0134219P.
XX PR 14-MAY-1999; 99US-0134221P.
XX PR 14-MAY-1999; 99US-0134370P.
XX PR 18-MAY-1999; 99US-0134768P.
```

PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145813P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 23-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 25-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151067P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159291P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.

```
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161933P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      68.9%; Score 45.5; DB 3; Length 252;
Best Local Similarity 77.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CY-WGCGYW 8
Db 173 CYRWGCGW 181

RESULT 15
AAG38268
ID AAG38268 standard; protein; 252 AA.
XX AC AAG38268;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 47184.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 25-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127452P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 23-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.

PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
```

PR 05-AUG-1999; 99US-01471192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-01493368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 23-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161320P.
PR 28-OCT-1999; 99US-0161392P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-01621142P.

Query Match 68.9%; Score 45.5; DB 3; Length 252;
Best Local Similarity 77.8%; Pred. No. 2.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 CY-WCGGYW 8
Db 173 CYRMCGGW 181

Search completed: April 18, 2005, 19:50:37
Job time : 90.0465 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 19:52:14 ; Search time 58.6047 Seconds
(without alignments)
45.371 Million cell updates/sec

Title: SEQ5

Perfect score: 66

Sequence: 1 cywgcgyw 8

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1421835 seqs, 332370583 residues

Total number of hits satisfying chosen parameters: 1421835

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Published Applications AA.*

- 1: /cgn2_6/ptodata/1/pubpaa/US07_PUBCOMB.pep.*
- 2: /cgn2_6/ptodata/1/pubpaa/PCT_NEW_PUB.pep.*
- 3: /cgn2_6/ptodata/1/pubpaa/US06_NEW_PUB.pep.*
- 4: /cgn2_6/ptodata/1/pubpaa/US06_PUBCOMB.pep.*
- 5: /cgn2_6/ptodata/1/pubpaa/US07_NEW_PUB.pep.*
- 6: /cgn2_6/ptodata/1/pubpaa/PCTUS_PUBCOMB.pep.*
- 7: /cgn2_6/ptodata/1/pubpaa/US08_NEW_PUB.pep.*
- 8: /cgn2_6/ptodata/1/pubpaa/US08_PUBCOMB.pep.*
- 9: /cgn2_6/ptodata/1/pubpaa/US09A_PUBCOMB.pep.*
- 10: /cgn2_6/ptodata/1/pubpaa/US09B_PUBCOMB.pep.*
- 11: /cgn2_6/ptodata/1/pubpaa/US09C_PUBCOMB.pep.*
- 12: /cgn2_6/ptodata/1/pubpaa/US09_NEW_PUB.pep.*
- 13: /cgn2_6/ptodata/1/pubpaa/US10A_PUBCOMB.pep.*
- 14: /cgn2_6/ptodata/1/pubpaa/US10B_PUBCOMB.pep.*
- 15: /cgn2_6/ptodata/1/pubpaa/US10C_PUBCOMB.pep.*
- 16: /cgn2_6/ptodata/1/pubpaa/US10D_PUBCOMB.pep.*
- 17: /cgn2_6/ptodata/1/pubpaa/US10_NEW_PUB.pep.*
- 18: /cgn2_6/ptodata/1/pubpaa/US11_NEW_PUB.pep.*
- 19: /cgn2_6/ptodata/1/pubpaa/US60_NEW_PUB.pep.*
- 20: /cgn2_6/ptodata/1/pubpaa/US60_PUBCOMB.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	48	72.7	54	17	US-10-776-224-340 Sequence 340, App
2	46	69.7	14	16	US-10-742-379-33 Sequence 33, Appl
3	46	69.7	50	16	US-10-742-379-136 Sequence 136, App
4	46	69.7	265	15	US-10-282-122A-61125 Sequence 61125, A
5	46	69.7	283	16	US-10-767-701-42031 Sequence 42031, A
6	45.5	68.9	72	16	US-10-437-963-190474 Sequence 190474, A
7	45.5	68.9	217	15	US-10-424-599-214537 Sequence 214537, A
8	45.5	68.9	292	15	US-10-425-114-69268 Sequence 69268, A
9	45	68.2	444	15	US-10-369-493-21261 Sequence 21261, A
10	45	68.2	445	15	US-10-369-493-463 Sequence 463, App
11	44	66.7	304	16	US-10-437-963-155759 Sequence 155759, A
12	42	63.6	46	15	US-10-424-599-238640 Sequence 238640, A
13	42	63.6	345	14	US-10-106-698-6232 Sequence 6232, Ap

14	42	63.6	533	16	US-10-704-365-2	Sequence 2, Appli
15	42	63.6	533	16	US-10-755-889-308	Sequence 308, App
16	42	63.6	2972	8	US-08-424-550B-387	Sequence 387, App
17	41	62.1	32	9	US-09-864-761-44928	Sequence 44928, A
18	41	62.1	69	9	US-09-864-761-46183	Sequence 46183, A
19	41	62.1	89	16	US-10-437-963-136942	Sequence 136942, A
20	41	62.1	552	16	US-10-322-281-732	Sequence 732, App
21	41	62.1	608	13	US-10-029-009-11	Sequence 11, Appl
22	41	62.1	626	13	US-10-029-009-23	Sequence 23, Appl
23	41	62.1	1376	16	US-10-437-963-161629	Sequence 161629, A
24	40.5	61.4	1604	9	US-09-746-491-8	Sequence 8, Appli
25	40	60.6	17	14	US-10-281-478-129	Sequence 129, App
26	40	60.6	59	15	US-10-424-599-223183	Sequence 223183, A
27	40	60.6	338	15	US-10-369-493-16505	Sequence 16505, A
28	40	60.6	350	14	US-10-017-161-2294	Sequence 2294, Ap
29	40	60.6	350	15	US-10-292-798-1940	Sequence 1940, Ap
30	40	60.6	470	15	US-10-369-493-13063	Sequence 13063, A
31	40	60.6	511	15	US-10-369-493-3925	Sequence 3925, Ap
32	40	60.6	520	15	US-10-282-122A-67162	Sequence 67162, A
33	40	60.6	522	15	US-10-369-493-2418	Sequence 2418, Ap
34	40	60.6	523	9	US-09-767-007A-2	Sequence 2, Appli
35	40	60.6	524	9	US-09-870-759-60	Sequence 60, Appl
36	40	60.6	524	10	US-09-751-708A-60	Sequence 60, Appl
37	40	60.6	524	15	US-10-267-502-386	Sequence 386, App
38	40	60.6	524	16	US-10-408-765A-1207	Sequence 1207, Ap
39	40	60.6	524	16	US-10-746-442-23	Sequence 23, Appl
40	40	60.6	527	9	US-09-870-759-61	Sequence 61, Appl
41	40	60.6	527	10	US-09-751-708A-61	Sequence 61, Appl
42	40	60.6	527	14	US-10-060-036-73	Sequence 73, Appl
43	40	60.6	697	16	US-10-322-281-729	Sequence 729, App
44	40	60.6	2280	16	US-10-437-963-198400	Sequence 198400, A
45	39	59.1	12	16	US-10-692-151-12	Sequence 12, Appl

ALIGNMENTS

RESULT 1

US-10-776-224-340
; Sequence 340, Application US/10776224
; Publication No. US20050074849A1
; GENERAL INFORMATION:
; APPLICANT: Gaudernack, Gustav
; APPLICANT: Ericksen, Jon Amund
; APPLICANT: Moller, Mona
; APPLICANT: Giertsen, Marianne Klemp
; APPLICANT: Saeterdal, Ingvil
; TITLE OF INVENTION: Peptides
; FILE REFERENCE: 01702.4015LO
; CURRENT APPLICATION NUMBER: US/10/776,224
; CURRENT FILING DATE: 2004-02-12
; PRIOR APPLICATION NUMBER: US 09/674,973
; PRIOR FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 459
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 340
; LENGTH: 54
; TYPE: PPT
; ORGANISM: Homo sapiens
US-10-776-224-340

Query Match 72.7%; Score 48; DB 17; Length 54;

Best Local Similarity 75.0%; Pred. No. 19;

Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 CYWGCYW 8

Db 5 CEWGCWS 12

RESULT 2

US-10-742-379-33

; Sequence 33, Application US/10742379

Publication No. US20040181033A1

GENERAL INFORMATION:
 APPLICANT: Han, HQ
 APPLICANT: Min, Hosung
 APPLICANT: Boone, Thomas Charles
 TITLE OF INVENTION: BINDING AGENTS WHICH INHIBIT MYOSTATIN
 FILE REFERENCE: A-828 (US)
 CURRENT APPLICATION NUMBER: US/10/742,379
 CURRENT FILING DATE: 2003-12-19
 PRIOR APPLICATION NUMBER: US 60/435,923
 PRIOR FILING DATE: 2002-12-20
 NUMBER OF SEQ ID NOS: 634
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 33
 LENGTH: 14
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Myostatin Binding Peptide
 US-10-742-379-33

Query Match 69.7%; Score 46; DB 16; Length 14;
 Best Local Similarity 71.4%; Pred. No. 13;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 YWCGGYW 8
 :|||
 Db 1 HWACGYW 7

RESULT 3

US-10-742-379-136
 Sequence 136, Application US/10742379
 Publication No. US20040181033A1

GENERAL INFORMATION:
 APPLICANT: Han, HQ
 APPLICANT: Min, Hosung
 APPLICANT: Boone, Thomas Charles
 TITLE OF INVENTION: BINDING AGENTS WHICH INHIBIT MYOSTATIN
 FILE REFERENCE: A-828 (US)
 CURRENT APPLICATION NUMBER: US/10/742,379
 CURRENT FILING DATE: 2003-12-19
 PRIOR APPLICATION NUMBER: US 60/435,923
 PRIOR FILING DATE: 2002-12-20
 NUMBER OF SEQ ID NOS: 634
 SOFTWARE: PatentIn version 3.2
 SEQ ID NO 136
 LENGTH: 50
 TYPE: PRT
 ORGANISM: Artificial Sequence
 FEATURE:
 OTHER INFORMATION: Myostatin Binding Peptide
 US-10-742-379-136

Query Match 69.7%; Score 46; DB 16; Length 50;
 Best Local Similarity 71.4%; Pred. No. 32;
 Matches 5; Conservative 1; Mismatches 1; Indels 0; Gaps 0;

QY 2 YWCGGYW 8
 :|||
 Db 1 HWACGYW 7

RESULT 4

US-10-282-122A-61125
 Sequence 61125, Application US/10282122A
 Publication No. US20040029129A1

GENERAL INFORMATION:
 APPLICANT: Wang, Liangsu
 APPLICANT: Zamudio, Carlos
 APPLICANT: Malone, Cheryl
 APPLICANT: Haselbeck, Robert
 APPLICANT: Ohlsen, Karl

APPLICANT: Zyskind, Judith
 APPLICANT: Wall, Daniel
 APPLICANT: Trawick, John
 APPLICANT: Carr, Grant
 APPLICANT: Yamamoto, Robert
 APPLICANT: Forsyth, R.
 APPLICANT: Xu, H.
 TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
 FILE REFERENCE: ELITRA 034A
 CURRENT APPLICATION NUMBER: US/10/282,122A
 CURRENT FILING DATE: 2003-02-20
 PRIOR APPLICATION NUMBER: 60/191,078
 PRIOR FILING DATE: 2000-03-21
 PRIOR APPLICATION NUMBER: 60/206,848
 PRIOR FILING DATE: 2000-05-23
 PRIOR APPLICATION NUMBER: 60/207,727
 PRIOR FILING DATE: 2000-05-26
 PRIOR APPLICATION NUMBER: 60/230,335
 PRIOR FILING DATE: 2000-09-06
 PRIOR APPLICATION NUMBER: 60/230,347
 PRIOR FILING DATE: 2000-09-09
 PRIOR APPLICATION NUMBER: 60/242,578
 PRIOR FILING DATE: 2000-10-23
 PRIOR APPLICATION NUMBER: 60/253,625
 PRIOR FILING DATE: 2000-11-27
 PRIOR APPLICATION NUMBER: 60/257,931
 PRIOR FILING DATE: 2000-12-22
 PRIOR APPLICATION NUMBER: 60/267,636
 PRIOR FILING DATE: 2001-02-09
 PRIOR APPLICATION NUMBER: 60/269,308
 PRIOR FILING DATE: 2001-02-16
 Remaining Prior Application data removed - See File Wrapper or PALM.
 NUMBER OF SEQ ID NOS: 78614
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 61125
 LENGTH: 265
 TYPE: PRT
 ORGANISM: Legionella pneumophila
 US-10-282-122A-61125

Query Match 69.7%; Score 46; DB 15; Length 265;
 Best Local Similarity 62.5%; Pred. No. 1.1e+02;
 Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CYWCGGYW 8
 :|||
 Db 65 CFWACGYF 72

RESULT 5

US-10-767-701-42031
 Sequence 42031, Application US/10767701
 Publication No. US20040172684A1

GENERAL INFORMATION:
 APPLICANT: Kovalic, David K.
 APPLICANT: Zhou, Yihua
 APPLICANT: Cao, Yongwei
 TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
 TITLE OF INVENTION: Plants and Uses Thereof For Plant Improvement
 FILE REFERENCE: 38-21(53535)B
 CURRENT APPLICATION NUMBER: US/10/767,701
 CURRENT FILING DATE: 2004-01-29
 NUMBER OF SEQ ID NOS: 63128
 SEQ ID NO 42031
 LENGTH: 283
 TYPE: PRT
 ORGANISM: Sorghum bicolor
 FEATURE:
 OTHER INFORMATION: Clone ID: SORBI-28MAY03-C13148_1.pep
 US-10-767-701-42031

Query Match 69.7%; Score 46; DB 16; Length 283;
 Best Local Similarity 85.7%; Pred. No. 1.2e+02;

Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 CYWGCY 7
| | | | |
Db 1 CKWGCY 7

RESULT 6

US-10-437-963-190474
; Sequence 190474, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 190474
; LENGTH: 72
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_86886C.1.pep
US-10-437-963-190474

Query Match 68.9%; Score 45.5; DB 16; Length 72;

Best Local Similarity 63.6%; Pred. No. 49;
Matches 7; Conservative 0; Mismatches 1; Indels 3; Gaps 1;

QY 1 CYWGC---GYW 8
| | | | |
Db 55 CRWGCCHRGYW 55

RESULT 7

US-10-424-599-214537
; Sequence 214537, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 214537
; LENGTH: 217
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_35754C.1.pep
US-10-424-599-214537

Query Match 68.9%; Score 45.5; DB 15; Length 217;

Best Local Similarity 77.8%; Pred. No. 11e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CY-WGCGY 8
| | | | |
Db 138 CYRWGCGGW 146

RESULT 8

US-10-425-114-69268
; Sequence 69268, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53113)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 69268
; LENGTH: 292
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: JC-GMFL02220141E06_FLI.pep
US-10-425-114-69268

Query Match 68.9%; Score 45.5; DB 15; Length 292;

Best Local Similarity 77.8%; Pred. No. 1.4e+02;
Matches 7; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 CY-WGCGY 8
| | | | |
Db 213 CYRWGCGGW 221

RESULT 9

US-10-369-493-21261
; Sequence 21261, Application US/10369493
; Publication No. US20030233675A1
; GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; APPLICANT: Hinkle, Gregory J.
; APPLICANT: Slater, Steven C.
; APPLICANT: Goldman, Barry S.
; APPLICANT: Chen, Xianfeng
; TITLE OF INVENTION: EXPRESSION OF MICROBIAL PROTEINS IN PLANTS FOR PRODUCTION OF
; FILE REFERENCE: 38-10(52052)B
; CURRENT APPLICATION NUMBER: US/10/369,493
; CURRENT FILING DATE: 2003-02-28
; PRIOR APPLICATION NUMBER: US 60/360,039
; PRIOR FILING DATE: 2002-02-21
; NUMBER OF SEQ ID NOS: 47374
; SEQ ID NO 21261
; LENGTH: 444
; TYPE: PRT
; ORGANISM: Xenorhabdus nematophilus
US-10-369-493-21261

Query Match 68.2%; Score 45; DB 15; Length 444;

Best Local Similarity 63.6%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 2 YWGC---GYW 8
| | | | |
Db 84 YWGLTAWGYW 94

RESULT 10

US-10-369-493-463
; Sequence 463, Application US/10369493
; Publication No. US20030233675A1

GENERAL INFORMATION:
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 238640
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57517C.1.pep
US-10-424-599-238640

Query Match 68.2%; Score 45; DB 15; Length 445;
Best Local Similarity 63.6%; Pred. No. 2.2e+02;
Matches 7; Conservative 0; Mismatches 0; Indels 4; Gaps 1;

QY 2 YWGC-----GYW 8
|||||
Db 80 YWGLTAWGYW 90

RESULT 11
US-10-437-963-155759
; Sequence 155759, Application US/10437963
; Publication No. US20040323343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 155759
; LENGTH: 304
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_55493C.1.pep
US-10-437-963-155759

Query Match 66.7%; Score 44; DB 16; Length 304;
Best Local Similarity 62.5%; Pred. No. 2.2e+02;
Matches 5; Conservative 2; Mismatches 1; Indels 0; Gaps 0;

QY 1 CYWGCYGW 8
|:|||||
Db 275 CFYGCNWN 282

RESULT 12
US-10-424-599-238640
; Sequence 238640, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou Yihua

APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 238640
; LENGTH: 46
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_57517C.1.pep
US-10-424-599-238640

Query Match 63.6%; Score 42; DB 15; Length 46;
Best Local Similarity 100.0%; Pred. No. 1e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CYWGC 5
|||||
Db 38 CFWGC 42

RESULT 13
US-10-106-698-6232
; Sequence 6232, Application US/10106698
; Publication No. US20030109690A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Colon and Colon Cancer Associated Polynucleotides and Polypeptide
; FILE REFERENCE: PA005P1
; CURRENT APPLICATION NUMBER: US/10/106,698
; CURRENT FILING DATE: 2002-03-27
; PRIOR APPLICATION NUMBER: PCT/US00/26524
; PRIOR FILING DATE: 2000-09-28
; PRIOR APPLICATION NUMBER: US 60/157,137
; PRIOR FILING DATE: 1999-09-29
; PRIOR APPLICATION NUMBER: US 60/163,280
; PRIOR FILING DATE: 1999-11-03
; NUMBER OF SEQ ID NOS: 8564
; SOFTWARE: Patent In Ver. 3.0
; SEQ ID NO 6232
; LENGTH: 345
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-106-698-6232

Query Match 63.6%; Score 42; DB 14; Length 345;
Best Local Similarity 100.0%; Pred. No. 4.5e+02;
Matches 5; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 CYWGC 5
|||||
Db 111 CYWGC 115

RESULT 14
US-10-704-365-2
; Sequence 2, Application US/10704365
; Publication No. US20040131605A1
; GENERAL INFORMATION:
; APPLICANT: Qin et al.
; TITLE OF INVENTION: PRECURSOR OF N-ACETYLGALACTOSAMINE-4-SULFATASE, METHODS OF TREAT
; FILE REFERENCE: 30610/30011A
; CURRENT APPLICATION NUMBER: US/10/704,365
; CURRENT FILING DATE: 2003-11-07
; PRIOR APPLICATION NUMBER: US 10/290,908
; PRIOR FILING DATE: 2002-11-07
; PRIOR APPLICATION NUMBER: US 10/317,249
; PRIOR FILING DATE: 2002-12-10
; PRIOR APPLICATION NUMBER: To be assigned

Tue Apr 19 06:58:40 2005

```
; PRIOR FILING DATE: 2003-09-09
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 2
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-704-365-2

Query Match      63.6%; Score 42; DB 16; Length 533;
Best Local Similarity 85.7%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 YMGCGYW 8
Db      444 YPGCGYW 450

RESULT 15
US-10-755-889-308
; Sequence 308, Application US/10755889
; Publication No. US20040171823A1
; GENERAL INFORMATION:
; APPLICANT: Bristol-Myers Squibb Company
; TITLE OF INVENTION: POLYNUCLEOTIDES AND POLYPEPTIDES ASSOCIATED WITH THE NF-KB
; FILE OF INVENTION: PATHWAY
; FILE REFERENCE: D0284 NP
; CURRENT APPLICATION NUMBER: US/10/755,889
; CURRENT FILING DATE: 2004-01-13
; PRIOR APPLICATION NUMBER: U.S. 60/440,068
; PRIOR FILING DATE: 2003-01-14
; PRIOR APPLICATION NUMBER: U.S. 60/469,757
; PRIOR FILING DATE: 2003-05-12
; NUMBER OF SEQ ID NOS: 823
; SOFTWARE: PatentIn version 3.2
; SEQ ID NO 308
; LENGTH: 533
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-755-889-308

Query Match      63.6%; Score 42; DB 16; Length 533;
Best Local Similarity 85.7%; Pred. No. 6.2e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY      2 YMGCGYW 8
Db      444 YPGCGYW 450
```

Search completed: April 18, 2005, 20:29:28
Job time : 59.6047 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM protein - protein search, using sw model

Run on: April 18, 2005, 19:38:48 ; Search time 22.3256 Seconds
(without alignments)
26.749 Million cell updates/sec

Title: SEQ5
Perfect score: 66
Sequence: 1 cywgcgyw 8

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 513545 seqs, 74649064 residues

Total number of hits satisfying chosen parameters: 513545

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5A COMB.pep.*
2: /cgn2_6/ptodata/1/iaa/5B COMB.pep.*
3: /cgn2_6/ptodata/1/iaa/6A COMB.pep.*
4: /cgn2_6/ptodata/1/iaa/6B COMB.pep.*
5: /cgn2_6/ptodata/1/iaa/PCTUS COMB.pep.*
6: /cgn2_6/ptodata/1/iaa/backfiles1.pep.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48	72.7	54	US-09-674-973A-340	Sequence 340, App
2	42	63.6	533	US-08-445-586-10	Sequence 10, Appl
3	42	63.6	533	US-08-484-493-13	Sequence 13, Appl
4	42	63.6	533	US-08-484-493-13	Sequence 13, Appl
5	42	63.6	533	US-08-484-493-13	Sequence 13, Appl
6	42	63.6	533	US-08-484-493-13	Sequence 13, Appl
7	42	63.6	533	US-08-484-493-13	Sequence 13, Appl
8	42	63.6	2972	US-08-469-260A-387	Sequence 387, App
9	42	63.6	2972	US-08-488-446-387	Sequence 387, App
10	42	63.6	2972	US-08-467-344A-387	Sequence 387, App
11	42	63.6	2972	US-08-424-550B-387	Sequence 387, App
12	41	62.1	63	US-09-583-110-4873	Sequence 4873, Ap
13	41	62.1	71	US-09-107-433-4664	Sequence 4664, Ap
14	40	60.6	523	US-08-100-247-2	Sequence 2, Appl
15	40	60.6	523	US-08-483-146A-2	Sequence 2, Appl
16	40	60.6	523	US-08-232-513A-3	Sequence 3, Appl
17	40	60.6	523	US-08-484-594A-2	Sequence 2, Appl
18	40	60.6	523	US-09-076-258A-2	Sequence 2, Appl
19	40	60.6	524	US-08-756-031-2	Sequence 2, Appl
20	40	60.6	524	US-09-352-548-1	Sequence 1, Appl
21	40	60.6	524	US-09-949-016-6272	Sequence 6272, Ap
22	40	60.6	535	US-09-949-016-8603	Sequence 8603, Ap
23	39	59.1	12	US-09-419-381-12	Sequence 12, Appl
24	39	59.1	30	US-08-753-829A-9	Sequence 9, Appl
25	39	59.1	346	US-09-252-991A-21708	Sequence 21708, A
26	39	59.1	394	US-09-902-540-15415	Sequence 15415, A
27	39	59.1	409	US-09-310-363C-4	Sequence 4, Appl

28 39 59.1 409 4 US-10-117-015-4 Sequence 4, Appli
29 39 59.1 749 4 US-09-489-039A-12279 Sequence 12279, A
30 38.5 58.3 469 2 US-08-484-126-1 Sequence 1, Appli
31 38.5 58.3 469 4 US-09-374-909-1 Sequence 1, Appli
32 38.5 58.3 632 4 US-09-315-127-2 Sequence 2, Appli
33 38.5 58.3 632 4 US-09-315-127-3 Sequence 3, Appli
34 38.5 58.3 665 4 US-09-309-572-14 Sequence 14, Appl
35 38.5 58.3 665 4 US-09-718-096-14 Sequence 26, Appl
36 38.5 58.3 1312 4 US-09-554-572-26 Sequence 4299, Ap
37 38 57.6 58 4 US-09-621-976-4299 Sequence 5992, Ap
38 38 57.6 87 4 US-09-621-976-5992 Sequence 7, Appli
39 38 57.6 111 4 US-09-730-212C-7 Sequence 8, Appli
40 38 57.6 164 4 US-09-898-751A-8 Sequence 29457, A
41 38 57.6 383 4 US-09-252-991A-29457 Sequence 339, App
42 37.5 56.8 55 4 US-09-674-973A-339 Sequence 26841, A
43 37.5 56.8 339 4 US-09-252-991A-26841 Sequence 4, Appli
44 37 56.1 84 4 US-09-599-632-4 Sequence 18867, A
45 37 56.1 154 4 US-09-252-991A-18867

ALIGNMENTS

RESULT 1

US-09-674-973A-340
; Sequence 340, Application US/09674973A
; Patent No. 6759046
; GENERAL INFORMATION:
; APPLICANT: No. 6759046sk Hydro ASA
; TITLE OF INVENTION: Peptides
; FILE REFERENCE: 26625-296
; CURRENT APPLICATION NUMBER: US/09/674,973A
; CURRENT FILING DATE: 2001-06-04
; NUMBER OF SEQ ID NOS: 459
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 340
; LENGTH: 54
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-674-973A-340

Query Match 72.7%; Score 48; DB 4; Length 54;
Best Local Similarity 75.0%; Pred. NO. 4;
Matches 6; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

Qy 1 CYWGCYW 8
Db 5 CEWGCYW 12

RESULT 2

US-08-445-586-10
; Sequence 10, Application US/08445586
; Patent No. 5627050
; GENERAL INFORMATION:
; APPLICANT: Takeshita, Sunao
; APPLICANT: Ito, Toshimi
; APPLICANT: Otawara-Ramamoto, Yoko
; APPLICANT: Amann, Egon
; TITLE OF INVENTION: Bone-Related Sulfatase-Like Protein and
; TITLE OF INVENTION: Process for Its Production
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Finnegan, Henderson, Farabow, Garrett &
; ADDRESSEE: Dunner
; STREET: 1300 I Street, N.W.
; CITY: Washington
; STATE: D.C.
; COUNTRY: USA
; ZIP: 20005-3314
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible

OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/445,586
FILING DATE:
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/111,887
FILING DATE: 26-AUG-1993
APPLICATION NUMBER: JP 230030/92
FILING DATE: 28-AUG-1992
PRIOR APPLICATION DATA:
APPLICATION NUMBER: JP 324034/92
FILING DATE: 03-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Forman, David S.
REGISTRATION NUMBER: 33,694
REFERENCE/DOCKET NUMBER: 02481.1322-00000
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-408-4000
TELEFAX: 202-408-4400
INFORMATION FOR SEQ ID NO: 10:
SEQUENCE CHARACTERISTICS:
LENGTH: 533 amino acids
TYPE: amino acid
TOPOLOGY: linear
MOLECULE TYPE: peptide
US-08-445-586-10

Query Match 63.6%; Score 42; DB 1; Length 533;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YWCGGYW 8
Db 444 YPGCGYW 450

RESULT 3
US-08-484-493-13
Sequence 13, Application US/08484493
Patent No. 5728381
GENERAL INFORMATION:
APPLICANT: Wilson, Peter J
APPLICANT: Morris, Charles P
APPLICANT: Anson, Donald S
APPLICANT: Occhiodoro, Teresa
APPLICANT: Bielicki, Julie
APPLICANT: Clements, Peter R
APPLICANT: Hopwood, John J
TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
IDURONATE 2-SULFATASE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,493
FILING DATE: 07-JUN-1995
CLASSIFICATION: 424
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 991,973
FILING DATE: 17-DEC-1992
ATTORNEY/AGENT INFORMATION:

NAME: Digiglio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8416Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 533 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
US-08-484-493-13
Query Match 63.6%; Score 42; DB 1; Length 533;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy 2 YWCGGYW 8
Db 444 YPGCGYW 450

RESULT 4
US-08-484-494-13
Sequence 13, Application US/08484494
Patent No. 5798239
GENERAL INFORMATION:
APPLICANT: Wilson, Peter J
APPLICANT: Morris, Charles P
APPLICANT: Anson, Donald S
APPLICANT: Occhiodoro, Teresa
APPLICANT: Bielicki, Julie
APPLICANT: Clements, Peter R
APPLICANT: Hopwood, John J
TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF
IDURONATE 2-SULFATASE
NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent In Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/484,494
FILING DATE: 07-JUN-1995
CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 991,973
FILING DATE: 17-DEC-1992
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 8416Z
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 533 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear

MOLECULE TYPE: protein
US-08-484-494-13

Query Match 63.6%; Score 42; DB 1; Length 533;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YMGCGYW 8
| | | | |
Db 444 YPGCGYW 450

RESULT 5

US-08-345-212-13
; Sequence 13, Application US/08345212

; Patent No. 5932211

; GENERAL INFORMATION:

; APPLICANT: Wilson, Peter J

; APPLICANT: Morris, Charles P

; APPLICANT: Anson, Donald S

; APPLICANT: Occhiodoro, Teresa

; APPLICANT: Bielicki, Julie

; APPLICANT: Clements, Peter R

; APPLICANT: Hopwood, John J

; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF

; IDURONATE 2-SULFATASE

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Scully, Scott, Murphy & Presser

; STREET: 400 Garden City Plaza

; CITY: Garden City

; STATE: New York

; COUNTRY: USA

; ZIP: 11530

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/08/345,212

; FILING DATE:

; CLASSIFICATION: 435

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 991,973

; FILING DATE: 17-DEC-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Digiglio, Frank S

; REGISTRATION NUMBER: 31,346

; REFERENCE/DOCKET NUMBER: 8416Z

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 516-742-4343

; TELEFAX: 516-742-4366

; TELEX: 230 901 SANS UR

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 533 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-08-345-212-13

Query Match 63.6%; Score 42; DB 2; Length 533;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YMGCGYW 8
| | | | |
Db 444 YPGCGYW 450

RESULT 6

TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF

IDURONATE 2-SULFATASE

US-09-249-003-13
; Sequence 13, Application US/09249003

; Patent No. 6153188

; GENERAL INFORMATION:

; APPLICANT: Wilson, Peter J

; APPLICANT: Morris, Charles P

; APPLICANT: Anson, Donald S

; APPLICANT: Occhiodoro, Teresa

; APPLICANT: Bielicki, Julie

; APPLICANT: Clements, Peter R

; APPLICANT: Hopwood, John J

; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF

; IDURONATE 2-SULFATASE

; NUMBER OF SEQUENCES: 15

; CORRESPONDENCE ADDRESS:

; ADDRESSEE: Scully, Scott, Murphy & Presser

; STREET: 400 Garden City Plaza

; CITY: Garden City

; STATE: New York

; COUNTRY: USA

; ZIP: 11530

; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk

; COMPUTER: IBM PC compatible

; OPERATING SYSTEM: PC-DOS/MS-DOS

; SOFTWARE: Patent In Release #1.0, Version #1.25

; CURRENT APPLICATION DATA:

; APPLICATION NUMBER: US/09/249,003

; FILING DATE:

; CLASSIFICATION:

; PRIOR APPLICATION DATA:

; APPLICATION NUMBER: 07/991,973

; FILING DATE: 17-DEC-1992

; ATTORNEY/AGENT INFORMATION:

; NAME: Digiglio, Frank S

; REGISTRATION NUMBER: 31,346

; REFERENCE/DOCKET NUMBER: 8416Z

; TELECOMMUNICATION INFORMATION:

; TELEPHONE: 516-742-4343

; TELEFAX: 516-742-4366

; TELEX: 230 901 SANS UR

; INFORMATION FOR SEQ ID NO: 13:

; SEQUENCE CHARACTERISTICS:

; LENGTH: 533 amino acids

; TYPE: amino acid

; STRANDEDNESS: single

; TOPOLOGY: linear

; MOLECULE TYPE: protein

US-09-249-003-13

Query Match 63.6%; Score 42; DB 3; Length 533;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YMGCGYW 8
| | | | |
Db 444 YPGCGYW 450

RESULT 7

US-09-685-844-13

; Sequence 13, Application US/09685844

; Patent No. 6541254

; GENERAL INFORMATION:

; APPLICANT: Wilson, Peter J

; APPLICANT: Morris, Charles P

; APPLICANT: Anson, Donald S

; APPLICANT: Occhiodoro, Teresa

; APPLICANT: Bielicki, Julie

; APPLICANT: Clements, Peter R

; APPLICANT: Hopwood, John J

; TITLE OF INVENTION: GLYCOSYLATION VARIANTS OF

IDURONATE 2-SULFATASE

NUMBER OF SEQUENCES: 15
CORRESPONDENCE ADDRESS:
ADDRESSEE: Scully, Scott, Murphy & Presser
STREET: 400 Garden City Plaza
CITY: Garden City
STATE: New York
COUNTRY: USA
ZIP: 11530
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/685,844
FILING DATE: 10-OCT-2000
CLASSIFICATION: <Unknown>
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 07/991,973
ATTORNEY/AGENT INFORMATION:
NAME: Digiglio, Frank S
REGISTRATION NUMBER: 31,346
REFERENCE/DOCKET NUMBER: 84162
TELECOMMUNICATION INFORMATION:
TELEPHONE: 516-742-4343
TELEFAX: 516-742-4366
TELEX: 230 901 SANS UR
INFORMATION FOR SEQ ID NO: 13:
SEQUENCE CHARACTERISTICS:
LENGTH: 533 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: protein
SEQUENCE DESCRIPTION: SEQ ID NO: 13:
US-09-685-844-13

Query Match 63.6%; Score 42; DB 4; Length 533;
Best Local Similarity 85.7%; Pred. No. 1.7e+02;
Matches 6; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 2 YMGCGYW 8
| | | | |
Db 444 YPGCGYW 450

RESULT 8
US-08-469-260A-387
; Sequence 387, Application US/08469260A
; Patent No. 6451578
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUEHROFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:

; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/469,260A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: POREMBSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 387:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2972 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; US-08-469-260A-387

Query Match 63.6%; Score 42; DB 4; Length 2972;
Best Local Similarity 62.5%; Pred. No. 7.1e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CYWGGYW 8
| | | | |
Db 295 CQMSAYW 302

RESULT 9
US-08-488-446-387
; Sequence 387, Application US/08488446
; Patent No. 6558898
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUEHROFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; TITLE OF INVENTION: REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ABBOTT LABORATORIES D377/AP6D
; STREET: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/488,446
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:

```
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 387:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2972 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-446-387

Query Match 63.6%; Score 42; DB 4; Length 2972;
Best Local Similarity 62.5%; Pred. No. 7.1e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CYWGCGYW 8
Db 295 CQGSAYW 302

RESULT 10
US-08-467-344A-387
; Sequence 387, Application US/08467344A
; Patent No. 6586568
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUEHROFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESS: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/467,344A
; FILING DATE: 07-Jun-1995
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/424,550
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 387:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2972 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-488-446-387

Query Match 63.6%; Score 42; DB 4; Length 2972;
Best Local Similarity 62.5%; Pred. No. 7.1e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CYWGCGYW 8
Db 295 CQGSAYW 302

RESULT 11
US-08-424-550B-387
; Sequence 387, Application US/08424550B
; Patent No. 6720166
; GENERAL INFORMATION:
; APPLICANT: JOHN N. SIMONS
; APPLICANT: TAMI J. PILOT-MATIAS
; APPLICANT: GEORGE J. DAWSON
; APPLICANT: GEORGE G. SCHLAUDER
; APPLICANT: SURESH M. DESAI
; APPLICANT: THOMAS P. LEARY
; APPLICANT: ANTHONY SCOTT MUEHROFF
; APPLICANT: JAMES C. ERKER
; APPLICANT: SHERI L. BUIJK
; APPLICANT: ISA K. MUSHAWAR
; TITLE OF INVENTION: NON-A, NON-B, NON-C, NON-D, NON-E HEPATITIS
; REAGENTS AND METHODS FOR THEIR USE
; NUMBER OF SEQUENCES: 716
; CORRESPONDENCE ADDRESS:
; ADDRESS: 100 ABBOTT PARK ROAD
; CITY: ABBOTT PARK
; STATE: IL
; COUNTRY: USA
; ZIP: 60064-3500
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/424,550B
; FILING DATE:
; CLASSIFICATION: 435435
; ATTORNEY/AGENT INFORMATION:
; NAME: FOREMSKI, PRISCILLA E.
; REGISTRATION NUMBER: 33,207
; REFERENCE/DOCKET NUMBER: 5527.PC.01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 708-937-6365
; TELEFAX: 708-938-2623
; INFORMATION FOR SEQ ID NO: 387:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 2972 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
US-08-424-550B-387

Query Match 63.6%; Score 42; DB 4; Length 2972;
Best Local Similarity 62.5%; Pred. No. 7.1e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 1 CYWGCGYW 8
Db 295 CQGSAYW 302

RESULT 12
US-09-583-110-4873
; Sequence 4873, Application US/09583110
```

```
; Patent No. 6699703
; GENERAL INFORMATION:
; APPLICANT: Lynn Doucette-Stamm et al.
; TITLE OF INVENTION: Nucleic Acid and Amino Acid Sequences Relating to Streptococcus
; FILE REFERENCE: PATH00-07A
; CURRENT APPLICATION NUMBER: US/09/583,110
; PRIORITY FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: US 09/107,433
; PRIOR FILING DATE: 1998-06-30
; PRIOR APPLICATION NUMBER: US 60/085,131
; PRIOR FILING DATE: 1998-05-12
; PRIOR APPLICATION NUMBER: US 60/051,553
; PRIOR FILING DATE: 1997-07-02
; NUMBER OF SEQ ID NOS: 5322
; SEQ ID NO 4873
; LENGTH: 63
; TYPE: PRT
; ORGANISM: Streptococcus pneumoniae
US-09-583-110-4873

Query Match          62.1%; Score 41; DB 4; Length 63;
Best Local Similarity 50.0%; Pred. No. 40;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 CYWGCYGW 8
      |::|:|
Db      36 CFWGSFW 43

RESULT 13
US-09-107-433-4664
; Sequence 4664, Application US/09107433
; Patent No. 6800744
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID
; SEQUENCES RELATING TO STREPTOCOCCUS PNEUMONIAE FOR DIAGNO
; THERAPEUTICS
; NUMBER OF SEQUENCES: 5206
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM 1809660
; COMPUTER: <Unknown>
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: <Unknown>
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,433
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/ 085131
; FILING DATE: May 12, 1998
; APPLICATION NUMBER: 60/051553
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneka
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-011
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 4664:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 71 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
```

```
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Streptococcus pneumoniae
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...71
; SEQUENCE DESCRIPTION: SEQ ID NO: 4664:
US-09-107-433-4664

Query Match          62.1%; Score 41; DB 4; Length 71;
Best Local Similarity 50.0%; Pred. No. 45;
Matches 4; Conservative 2; Mismatches 2; Indels 0; Gaps 0;

QY      1 CYWGCYGW 8
      |::|:|
Db      44 CFWGSFW 51

RESULT 14
US-08-100-247-2
; Sequence 2, Application US/08100247
; Patent No. 5571787
; GENERAL INFORMATION:
; APPLICANT: O'BRIEN, JOHN S.
; APPLICANT: KISHIMOTO, YASUO
; TITLE OF INVENTION: PROSAPOSIN AS A NEUTROTROPHIC FACTOR
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: KNOBBE, MARTENS, OLSON AND BEAR
; STREET: 620 NEWPORT CENTER DRIVE SIXTEENTH FLOOR
; CITY: NEWPORT BEACH
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/100,247
; FILING DATE: 19930730
; CLASSIFICATION: 514
; ATTORNEY/AGENT INFORMATION:
; NAME: Israelsen, Ned A.
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: O'BRIEN.002A
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; HYPOTHETICAL: NO
; ANTI-SENSE: NO
; FRAGMENT TYPE: N-terminal
; IMMEDIATE SOURCE:
; CLONE: PROSAPOSIN
US-08-100-247-2

Query Match          60.6%; Score 40; DB 1; Length 523;
Best Local Similarity 62.5%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY      1 CYWGCYGW 8
      |::|:|
Db      494 CIWGPSYW 501
```

```

RESULT 15
US-08-483-146A-2
; Sequence 2, Application US/08483146A
; Patent No. 5696080
; GENERAL INFORMATION:
; APPLICANT: O'Brien, John S.
; APPLICANT: Kishimoto, Yasuo
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITIONS
; TITLE OF INVENTION: COMPRISING PROSAPOSIN AND NEUROTROPHIC PEPTIDES DERIVED
; TITLE OF INVENTION: THEREFROM
; NUMBER OF SEQUENCES: 11
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Knobbe, Martens, Olson and Bear
; STREET: 620 Newport Center Blvd. 16th Floor
; CITY: Newport Beach
; STATE: CA
; COUNTRY: USA
; ZIP: 92660
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/483,146A
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Israel, Ned A
; REGISTRATION NUMBER: 29,655
; REFERENCE/DOCKET NUMBER: MVELOS.002DV1
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-235-8550
; TELEFAX: 619-235-0176
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 523 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: peptide
; FRAGMENT TYPE: N-terminal
US-08-483-146A-2
Query Match 60.6%; Score 40; DB 1; Length 523;
Best Local Similarity 62.5%; Pred. No. 3.2e+02;
Matches 5; Conservative 0; Mismatches 3; Indels 0; Gaps 0;
QY 1 CYWGGYW 8
Db 494 CIWGPSYW 501

```

Search completed: April 18, 2005, 20:00:40
 Job time : 23.3256 secs

This Page Blank (uspto)